Sensor Network Design Proposal for the Battle of the Water Sensor Networks (BWSN)

Avi Ostfeld[1], Member ASCE, and Elad Salomons [2]

Abstract

This study presents a multiobjective solution approach to the Battle of the Water Sensor Networks (BWSN) initiative (Ostfeld et al., 2006). The developed methodology tailors the algorithm of Ostfeld and Salomons (2005) for optimally placing sensors in a water distribution system with the NSGA-II multiobjective genetic algorithm of Deb et al. (2002). Pareto optimal fronts are shown and discussed for the two BWSN Networks, for selected BWSN cases.

Introduction

During the last decade there has been an increasing interest in development of sensor networks to cope with both deliberate and accidental hazards intrusions into water distribution systems. Optimization models and solution algorithms have been developed for sensors locations using various algorithms and objectives. These optimization models have made simplifying assumptions about design objectives, network contaminant transport, sensor response, event detection, emergency response, installation and maintenance costs, etc. Little is known about how these design algorithms compare to the efforts of human designers, and thus what advantages they propose for practical design of sensor networks. To explore these issues the Battle of the Water Sensor Networks (BWSN) initiative was called upon (Ostfeld et al., 2006) with the purpose of objectively comparing the performance of contributed sensor network designs of different teams, as applied on two water distribution system examples. This manuscript summarizes such an effort.

Design objectives

Four design objectives for placing a set of sensors were defined (Ostfeld et al., 2006):

Expected Time of Detection (Z_1)

For a particular contamination scenario, the time of detection by a sensor is the elapsed time from the start of the contamination event to the first presence of a non-zero contaminant concentration at that sensor. Denote this time of detection t_i where the subscript i refers to the ith sensor location. The time of detection for the sensor *network* and this particular contamination event, t_d, is the minimum among all sensors present in the design, $t_d = \min_i t_i$.

---

[1] Senior Lecturer, Faculty of Civil and Environmental Engineering, Technion – Israel Institute of Technology, Haifa 32000, Israel; PH +972-4-8292782; FAX +972-4-8228898; email: ostfeld@tx.technion.ac.il

[2] Owner, OptiWater, 6 Amikam Israel St., Haifa, 34385, Israel; PH +972-54-2002050; FAX +972-15154-2002050; email: selad@optiwater.com

Water Distribution Systems Analysis Symposium 2006

The objective function to be minimized is the expected value computed over the assumed probability distribution of contamination events,

$$Z_1 = E(t_d),$$

where $E()$ denotes the mathematical expectation, and will be approximated by Monte-Carlo simulation for purposes of comparing contributed designs. The assumed probability distribution of contamination events is discussed below in conjunction with other design assumptions.

Expected Population Affected Prior to Detection (Z_2)

For a particular contamination scenario, the population affected is a function of the ingested contaminant mass. The ingested contaminant mass in turn depends on the time of detection for the sensor network, as described above; a key assumption is that no mass is ingested after detection. For a particular contamination scenario, the mass ingested – prior to detection – by any individual at network node i is calculated,

$$M_i = \phi \Delta t \sum_{k=1}^{N} c_{ik} \rho_{ik},$$

where ϕ is the mean volumetric ingestion rate (Liters/day), Δt is the evaluation time step (days), c_{ik} is the contaminant concentration for node i and time step k (mg/Liter), ρ_{ik} is a "dose rate multiplier" for node i and time step k (unitless), and N is the number of evaluation time steps prior to detection, *i.e.* the largest integer such that $N\Delta t \le t_d$. The series ρ_{ik}, $k=1,\ldots$ has a mean of 1 (so ϕ is truly the mean volumetric ingestion rate) and is intended to model the variation in ingestion rate throughout the day. We assume the ingestion rate varies with the water demand rate at the respective node, thus $\rho_{ik} = q_{ik} / \overline{q}_i$, where q_{ik} is the water demand for time step k and node i, and $\overline{q}_i$ is the average water demand at node i.

A typical dose-response model is used to express the probability that any person ingesting mass M_i will be affected (*i.e.* become infected or symptomatic),

$$R_i = \Phi[\beta \log_{10}((M_i / W) / D_{50})]$$

where R_i is the probability [0, 1] that a person who ingests contaminant mass M_i will become infected or symptomatic, β is the so-called Probit slope parameter (unitless), W is the assumed body weight (kg), D_{50} is the dose that would result in a 0.5 probability of becoming infected or symptomatic (mg/kg), and Φ is the Standard Normal Cumulative Distribution Function.

The population affected for a particular contamination scenario is calculated, $P_a = \sum_{i=1}^{m} R_i P_i$, where P_i is the population assigned to node i and m is the total number of nodes. The objective function to be minimized is the expected value of P_a computed over the assumed probability distribution of contamination events,

$$Z_2 = E(P_a),$$

where $E()$ denotes the mathematical expectation, and will be approximated by Monte-Carlo simulation for purposes of comparing contributed designs. The assumed probability distribution of contamination events is discussed below in conjunction with other design assumptions.

Water Distribution Systems Analysis Symposium 2006

Expected Demand of Contaminated Water Prior to Detection (Z_3)

Z_3 is the expected volume of consumed contaminated water prior to detection,

$$Z_3 = E\ (V_d)$$

where V_d denotes the total volumetric water demand that exceeds a predefined hazard concentration, and $E()$ is the mathematical expectation, which will be approximated by Monte-Carlo simulation for purposes of comparing contributed designs. As for the expected population affected, a key assumption is that no water is delivered after detection. The assumed probability distribution of contamination events is discussed below in conjunction with other design assumptions. Z_3 (as Z_2 and Z_1) is to be minimized.

Detection likelihood (Z_4)

Given a sensor network design (i.e., number and locations) the detection likelihood (Z_4 – the probability of detection) is defined:

$$Z_4 = \frac{1}{N} \sum_{i=1}^{N} d_i$$

where: d_i = 1 if contamination scenario i is detected, and zero otherwise, and N is the number of the total contamination scenarios considered. Z_4 is to be maximized.

Assumptions, cases, networks

Four cases (A-D) were defined (Ostfeld et al., 2006):

Base Case (A)

1. All quantities affecting network model water quality predictions are assumed to be known and deterministic. Sensor network designs will be challenged by an ensemble of contamination scenarios sampled from a statistical distribution; the probability distribution of contamination events is described as follows. Contaminant intrusions occur at network nodes, with injection flow rate of 125 Liter/hr, contaminant concentration of 230,000 mg/L, and injection duration of 2 hrs. The contaminant is stable after injection in finished drinking water. Each contamination scenario involves a single injection location, which may occur at any network node and begin at any time with equal probability. For purposes of design evaluation, contaminant concentrations will be predicted with a 5 minute time step.
2. For purposes of calculating the expected population affected prior to detection (Z_2): ϕ = 2 Liters/day, β = 0.34, D_{50} = 41 mg/kg, W = 70 kg. For purposes of estimating node population, the total per capita water consumption rate is assumed to be 300 Liters/day.
3. For purposes of calculating the expected demand of contaminated water prior to detection (Z_3) the hazard concentration threshold is C = 0.3 (mg/Liter).
4. Sensors instantly detect any non-zero contaminant concentration and action is taken to eliminate further exposure without delay.

Derivative Case (B)

Identical to base case (A) except that the injection duration is increased to 10 hrs.

Derivative Case (C)

Identical to base case (A) except that the response delay is 3 hour, *i.e.*, it takes 3 hr after detection for emergency response to limit contaminant exposure.

Derivative Case (D)

Identical to base case (A) except that any contamination scenario involves two injection locations, which may occur at any two distinct nodes with equal probability. The contamination scenario may begin at any time with equal probability, but both injections are synchronized to begin at the same time.

The four cases are summarized in Table 1. Figs. 1 and 2 show the two networks examples layouts.

Water Distribution Systems Analysis Symposium 2006

Table 1: Summary of design assumptions and cases

Base case (A)			**Derivative case**		
Component	**Characteristics**	**Description**	**B**	**C**	**D**
System	Possible injection locations	Nodes			
	Existing sensors	None			
	Demands	Deterministic			
Injection	Nodes injection probabilities	Even			
	Number of injection events	One (i.e., a single random attack)			Two
	Duration	2 (hr)	10 (hr)		
	Flow	125 (Liter/hr)			
	Constituent concentration	230000 (mg/Liter)			
	Constituent type	Conservative (i.e., no decay, no interaction)			
Sensors	Detection delay	Real time (i.e., zero delay)		3 (hr)	
	Detection sensitivity	Ideal (i.e., above zero)			
Design	Number of sensors	5 ; 20			

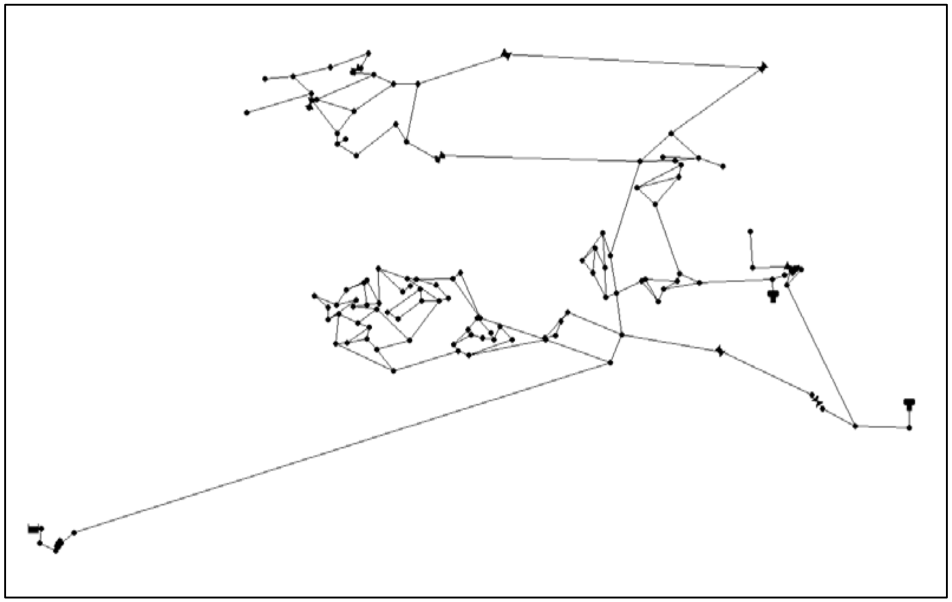

Fig. 1: Layout of Network 1 (126 nodes, 1 reservoir, 2 tanks, 168 pipes, 2 pumps, 8 valves) (USEPA, 2001)

Fig. 2: Layout of Network 2 (12523 nodes, 2 reservoirs, 2 tanks, 14822 pipes, 4 pumps, 5 valves) (USEPA, 2001)

Results

Although not written explicitly at the BWSN rules it was evident that there should be a clear tradeoff between Z_i (i = 1, 2, 3) and Z4, thus a unique "optimal" solution does not exist. As a result of that a multiobjective solution approach was adopted, tailoring the algorithm of Ostfeld and Salomons (2005) for optimally placing sensors in a water distribution system with the NSGA-II multiobjective genetic algorithm of Deb et al. (2002).

Network 1

Results for the outcome of Cases A-D are shown in Figs. 3 – 7, and at Table 2, respectively. For each of the Cases A-C a randomized matrix of size 12900 corresponding to 100 random injections at each of the system nodes, the reservoir, and the two tanks, was constructed. Those matrices were used to establish the Pareto optimal fronts shown in Figs. 3 – 5. For Case D a randomized matrix of size 15000 was used to create the Pareto optimal fronts shown in Fig. 6.

The following can be seen from Figs. 3 – 6: (1) A clear tradeoff between Z_i (i = 1, 2, 3) and Z4 for all cases. (2) Better Pareto optimal fronts using 20 sensors for all cases. (3) Compared to Case A, as the injection duration increases (Case B) so is the detection likelihood (i.e., Z4). (4) Compared to Case A, both the Z_i (i = 1, 2, 3) and the detection likelihood (Z4) values increase when two injection events are considered (Case D).

Fig. 7 shows the sensors layouts for two extreme solutions: 1 and 3 (see Fig. 3). It can be seen from Fig. 7 that for solution 1 which corresponds to the minimum of Z1 and Z4, the sensors are much more concentrated than for solution 3 in which both Z1 and Z4 are at their maximum values.

Network 2

Results for the outcome of Case A are shown in Figs. 8 and 9, and at Table 3, respectively.
To create the Pareto optimal fronts shown in Fig. 8 a randomized matrix of size 2000 was constructed.

It can be seen from Fig. 8 that: (1) A clear tradeoff between Z_i (i = 1, 2, 3) and Z4 exists. (2) Better Pareto optimal fronts were received using 20 sensors. (3) All Pareto fronts are inferior to the Pareto fronts of Network 1 Case A (see Fig. 3), which suggests that more sensors are needed for Network 2.

Fig. 9 shows the sensors layouts for solutions: 35 and 36 (see Fig. 8). It can be seen from Fig. 9 that at solution 35 the 5 sensors are relatively concentrated compared to solution 36 in which the 20 sensors are spread throughout the system.

Similar observations were received for cases B, C, and D.

Water Distribution Systems Analysis Symposium 2006

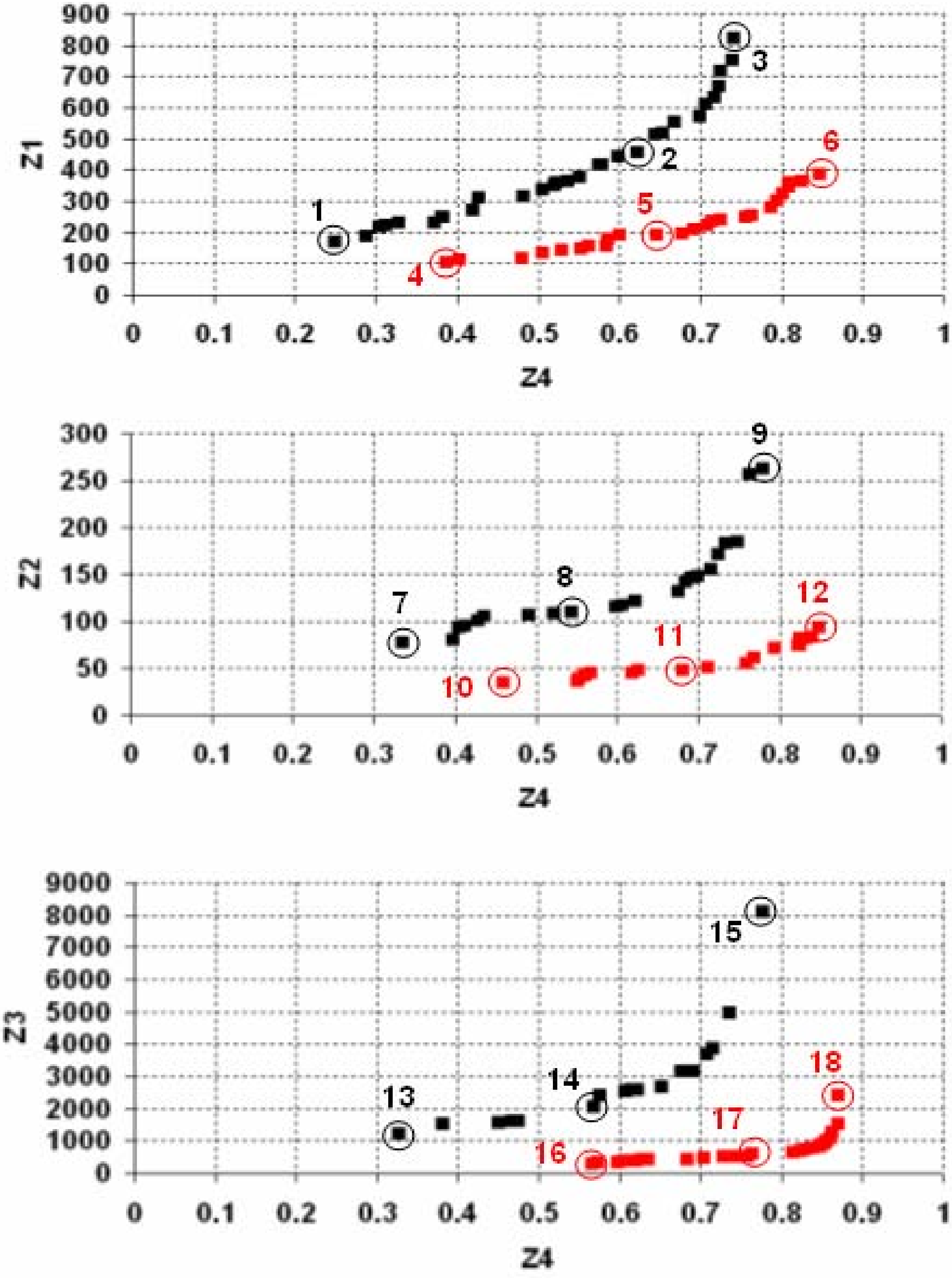

Fig. 3: Pareto optimal fronts for N1A5 (Network 1, Case A, five sensors ■) and for N1A20 (■)

Water Distribution Systems Analysis Symposium 2006

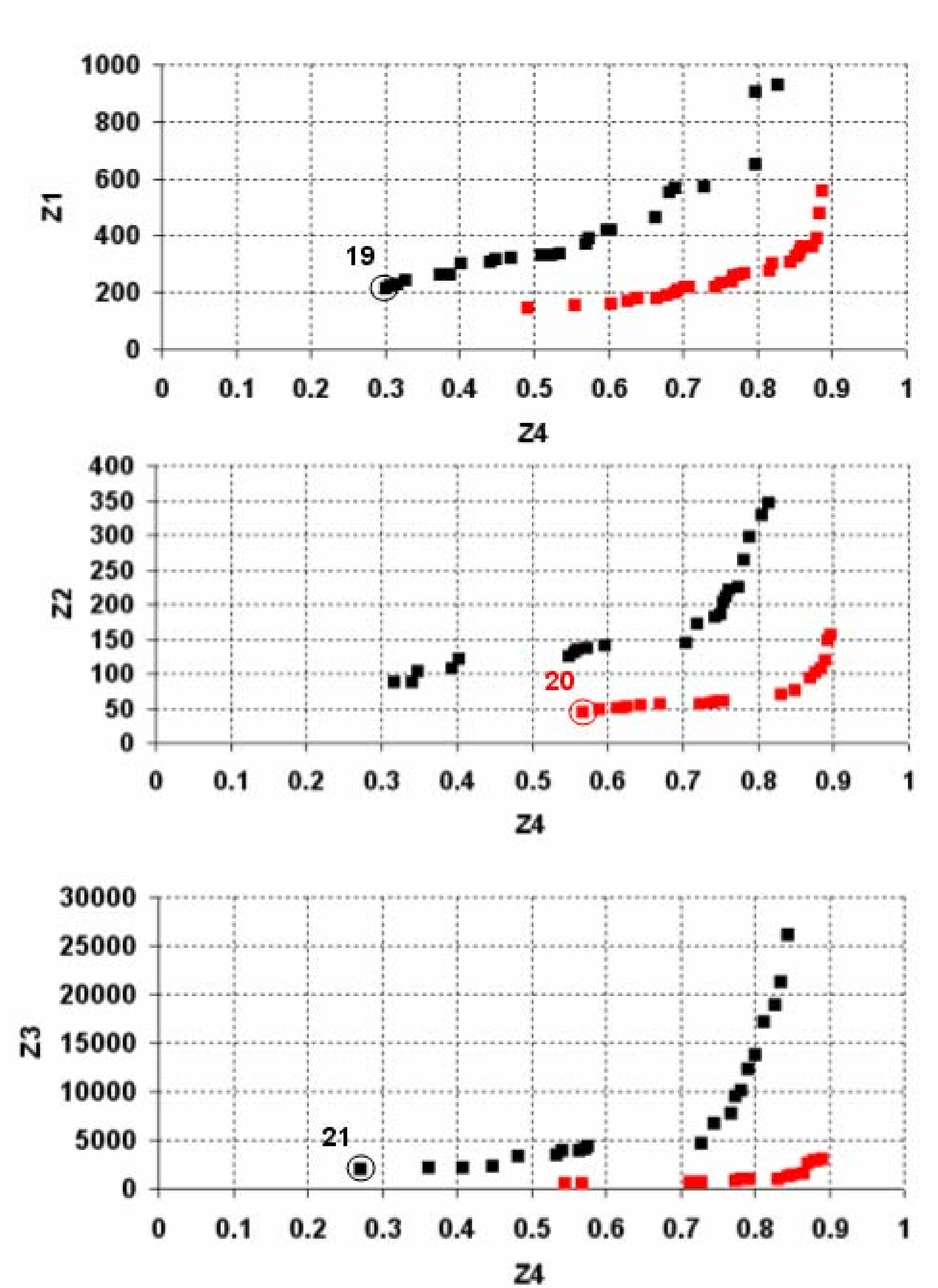

Fig. 4: Pareto optimal fronts for N1B5 (Network 1, Case B, five sensors ■) and for N1B20 (■)

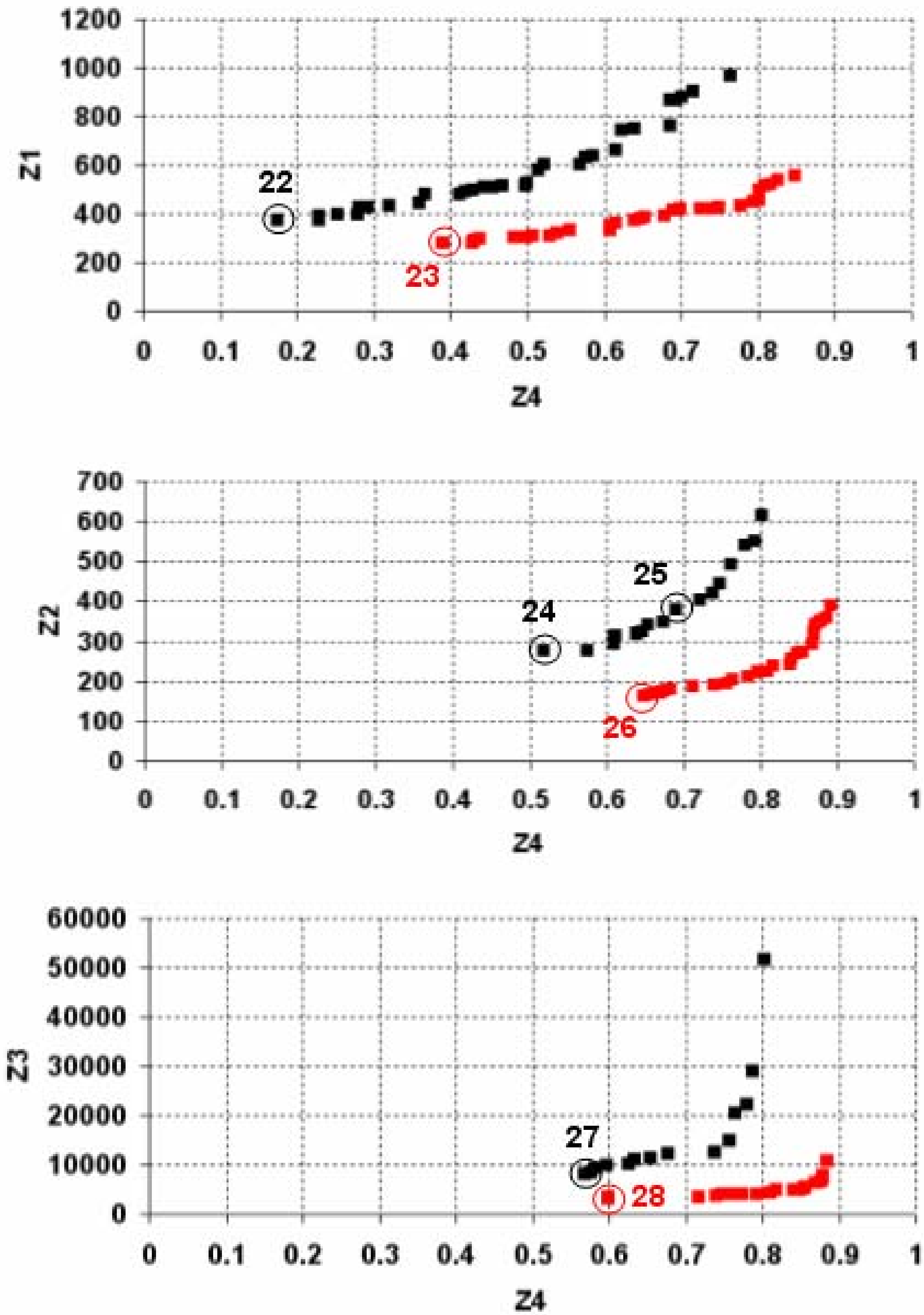

Fig. 5: Pareto optimal fronts for N1C5 (Network 1, Case C, five sensors ■) and for N1C20 (■)

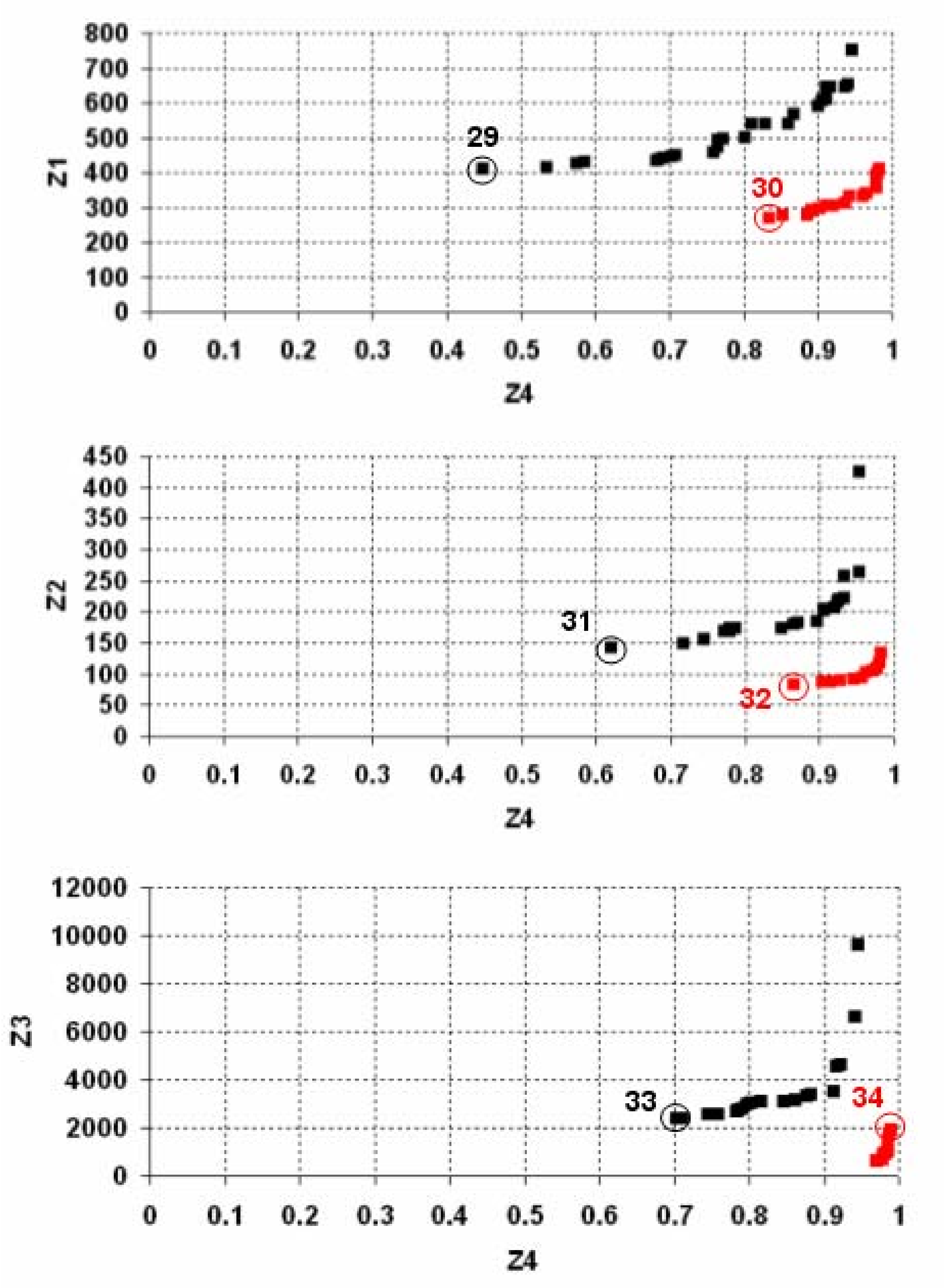

Fig. 6: Pareto optimal fronts for N1D5 (Network 1, Case D, five sensors ■) and for N1D20 (■)

Water Distribution Systems Analysis Symposium 2006

SOLUTION NUMBER (Z_i i = 1, 2, 3; Z4	SENSORS LOCATIONS (NETWORK JUNCTION NUMBERS)																			
1 (172 ; 0.25)	49	35	58	21	41															
2 (458 ; 0.62)	117	71	98	68	82															
3 (823 ; 0.74)	118	95	14	102	83															
4 (106 ; 0.39)	95	44	91	58	4	21	42	9	50	54	16	35	36	33	125	93	20	42	34	49
5 (195 ; 0.65)	68	5	40	65	51	69	88	89	22	72	34	71	53	112	63	78	122	28	118	97
6 (386 ; 0.85)	11	18	118	123	7	46	90	34	45	84	74	41	14	5	8	4	83	19	101	49
7 (78 ; 0.33)	65	116	49	30	54															
8 (111 ; 0.54)	96	116	68	81	22															
9 (263 ; 0.78)	100	117	68	83	45															
10 (36 ; 0.46)	40	30	54	71	65	37	76	53	72	27	58	94	65	22	69	73	59	21	97	51
11 (49 ; 0.68)	95	114	77	10	65	5	88	31	68	110	32	17	37	21	23	85	74	69	2	49
12 (94 ; 0.85)	11	100	118	99	17	46	81	37	124	82	26	122	37	5	31	68	83	19	86	109
13 (1246 ; 0.33)	110	37	68	20	48															
14 (2052 ; 0.57)	9	84	68	88	23															
15 (8136 ; 0.78)	10	80	118	122	83															
16 (326 ; 0.57)	41	88	5	30	26	21	65	7	17	31	36	11	90	1	10	124	87	34	116	37
17 (525 ; 0.76)	34	10	118	80	99	53	17	28	97	49	31	85	20	4	8	68	83	31	87	39
18 (2439 ; 0.87)	71	21	117	126	123	70	16	56	27	20	53	100	67	10	45	35	83	12	105	11
19 (215 ; 0.30)	58	25	9	37	21															
20 (46 ; 0.57)	68	49	94	40	73	15	93	9	115	62	28	92	3	31	37	56	10	34	93	75
21 (2044 ; 0.27)	115	44	39	31	110															
22 (374 ; 0.17)	23	40	50	35	49															
23 (282 ; 0.39)	102	45	94	42	58	120	97	29	38	24	122	106	39	55	56	35	34	48	109	52
24 (277 ; 0.52)	83	101	31	64	49															
25 (380 ; 0.69)	83	10	24	6	101															
26 (166 ; 0.65)	65	10	64	76	93	55	49	39	83	73	35	53	58	27	42	98	88	29	100	51
27 (8197 ; 0.57)	83	60	37	22	21															
28 (3672 ; 0.60)	60	124	71	93	99	49	85	37	28	23	34	78	43	41	28	68	83	81	100	52
29 (414 ; 0.45)	52	49	92	37	32															
30 (271 ; 0.83)	9	109	98	70	118	41	52	49	112	4	121	62	71	97	35	68	67	17	64	39
31 (142 ; 0.62)	37	95	17	31	48															
32 (84 ; 0.86)	102	117	17	54	96	21	62	104	24	122	37	89	12	98	30	6	9	115	68	70
33 (2430 ; 0.70)	37	68	102	49	17															
34 (1969 ; 0.99)	100	118	59	34	126	124	123	53	45	22	114	104	10	122	74	83	93	123	49	35

Table 2: Detailed sensor locations for Network 1 (see also Figs. 3 – 6)

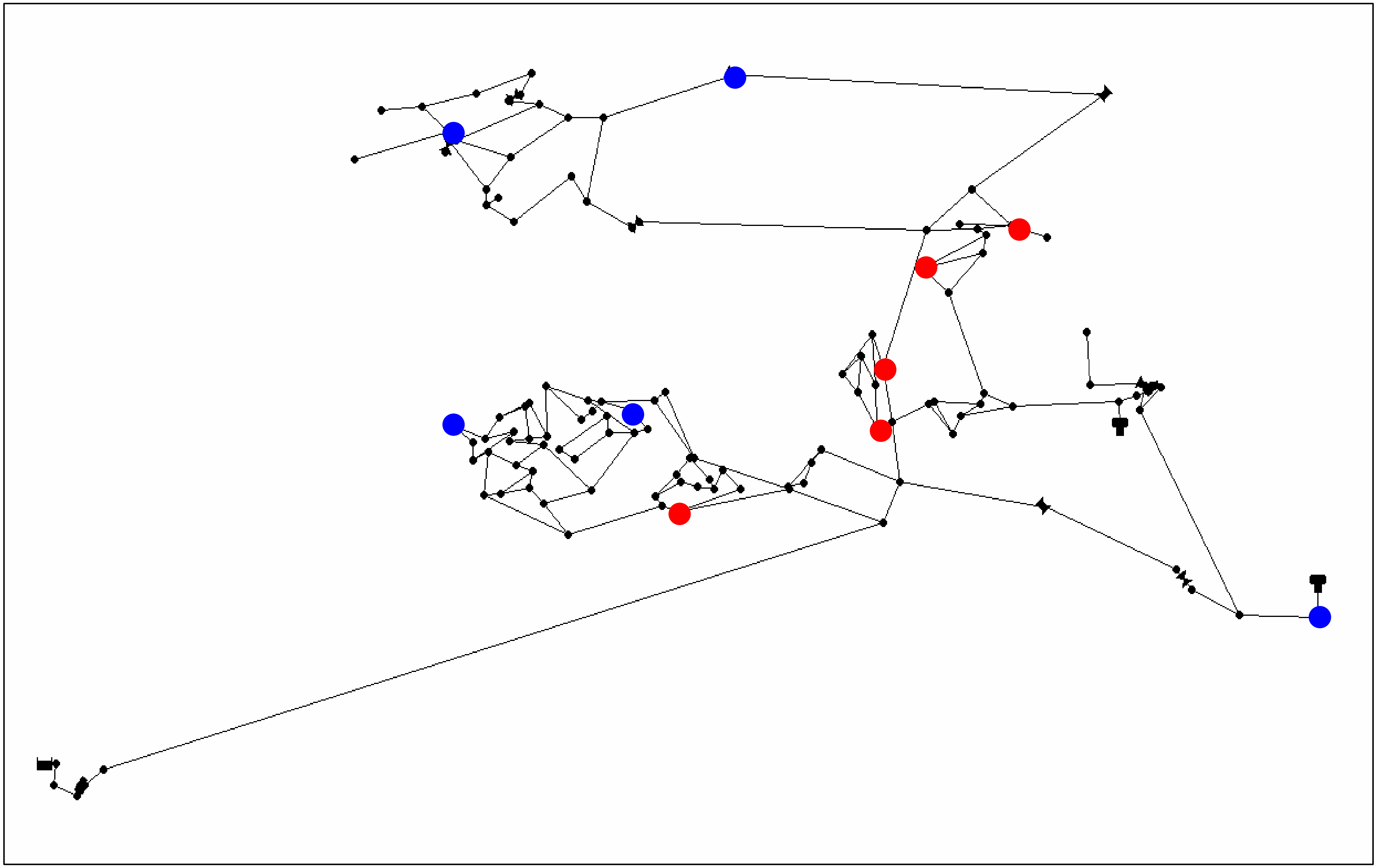

Fig. 7: Layout of sensors for Network 1 at points 1 (●) and 3 (●) (see Fig. 3, and Table 2)

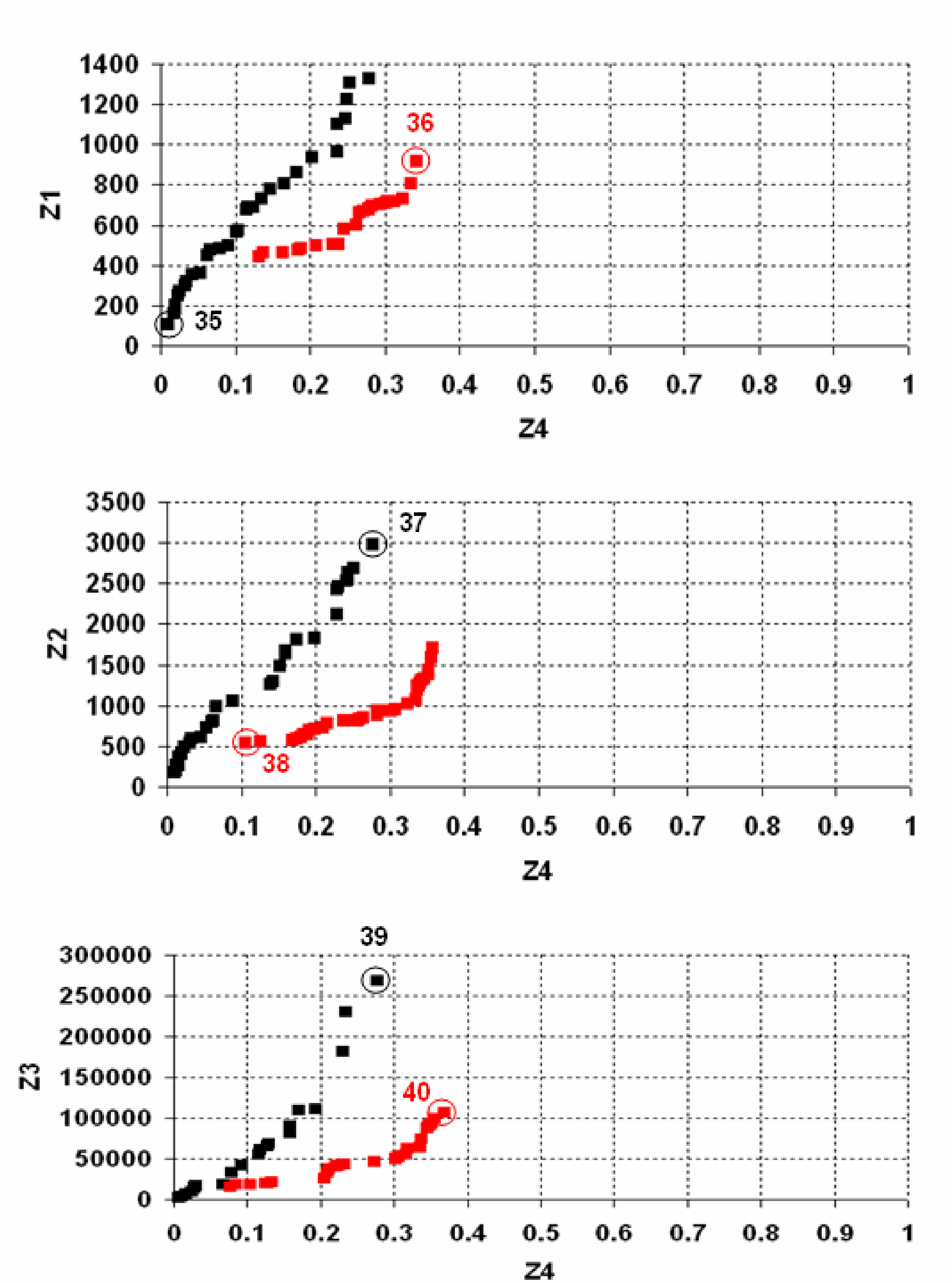

Fig. 8: Pareto optimal fronts for N2A5 (Network 2, Case A, five sensors ■) and for N2A20 (■)

Water Distribution Systems Analysis Symposium 2006

SOLUTION NUMBER (Z_i i = 1, 2, 3; Z4	SENSORS LOCATIONS (NETWORK JUNCTION NUMBERS)																			
35 (111 ; 0.01)	5810	6225	7291	12118	4409															
36 (917 ; 0.34)	11661	6022	7641	9909	8368	8599	4564	3042	6291	6306	10985	11105	3747	183	3640	9869	7714	11259	1696	9045
37 (2979 ; 0.28)	1958	4845	3713	2447	7906															
38 (557 ; 0.11)	8256	9426	8875	8804	7547	8534	134	11246	8243	11904	8217	8102	5053	3494	6993	5222	5905	3077	8376	3995
39 (269177 ; 0.28)	5039	4646	1515	3234	5541															
40 (107666 ; 0.37)	2872	4319	4782	3281	8766	3712	11184	4433	22	11623	8560	3129	9785	8098	10734	6738	7428	611	7669	7500

Table 3: Detailed sensor locations for Network 2 (see also Fig. 8)

Water Distribution Systems Analysis Symposium 2006

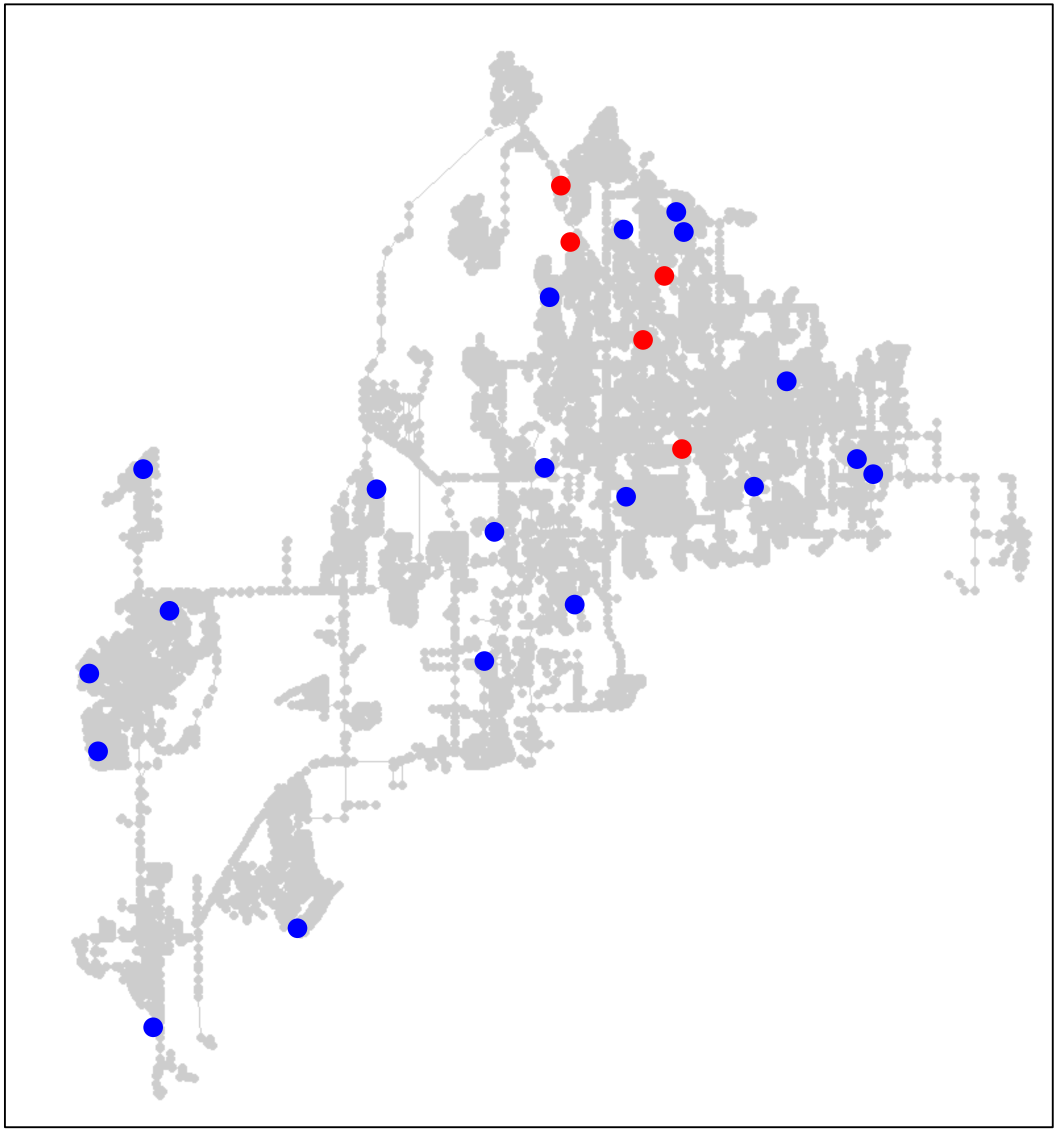

Fig. 9: Layout of sensors for Network 2 at points 35 (●) and 36 (●) (see Fig. 8, and Table 3)

Conclusions

This paper has presented results for the BWSN using a multiobjective approach. The results obtained could be explained from an engineering point of view in most cases, which gives raise to the conclusion that such an approach could also be used in reality. Still the main limitation is in sampling efficiently injection events in a manner which will satisfactorily approximate the objective function in hand. This is the major difficulty if the water distribution systems to be handled increase in size (e.g., Network 2 of the BWSN).

Acknowledgements

This work was funded by the Technion Grand Water Research Institute (GWRI), by NATO [Science for Peace (SfP) project no. CBD.MD.SFP 981456], and by the Technion LOEWENGART RESEARCH FUND.

References

Deb K., Pratap A., Agarwal S, and Meyarivan T. (2002). "A Fast and Elitist Multiobjective Genetic Algorithm: NSGA-II." IEEE Transactions on Evolutionary Computation, Vol. 6, No. 2, pp. 182-197.

Ostfeld A. and Salomons E. (2005). "Securing Water Distribution Systems Using Online Contamination Monitoring." Journal of Water Resources Planning and Management Division, ASCE, Vol. 131, No. 5, pp. 402-405.

Ostfeld A., Uber J., and Salomons S. (2006). "Battle of the Water Sensor Networks (BWSN): A Design Challenge for Engineers and Algorithms." Water Distribution Systems Analysis 2006 Cincinnati, Ohio U.S.A. August 27-29, 2006 (www.eng.uc.edu/wdsa2006/).

USEPA (2001). "EPANET 2.00.10" http://www.epa.gov/ORD/NRMRL/wswrd/epanet.htm

Water Distribution Systems Analysis Symposium 2006