Supplementary material

A novel hybrid entropy-clustering approach for optimal placement of pressure sensors for leakage detection in water distribution systems under uncertainty

Narges Taravatrooy, Mohammad Reza Nikoo, Salma Hobbi, Mojtaba Sadegh, Azizallah Izady

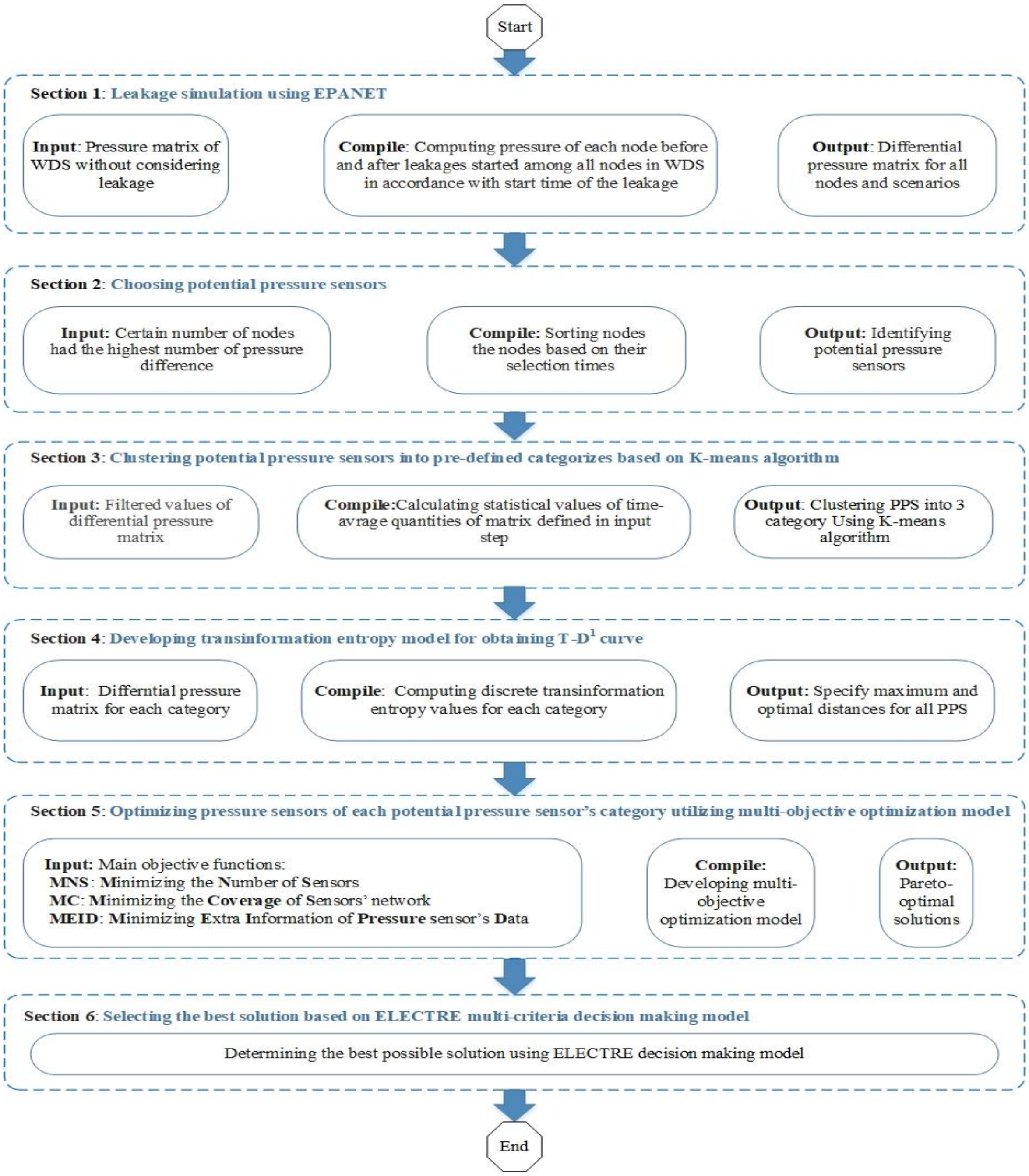


Fig. S1 Detailed information of the flowchart shown in Fig. 1

At the first section (including steps 1 to 4), leakage scenarios for all nodes of WDS are simulated by EPANET model (Rossman and others 2000). Firstly, pressure matrix of WDS system was established without any consideration of leakage. Then, different leakage scenarios for all nodes were evaluated according to C-Town's WDS data (Ostfeld et al. 2011) to determine matrices of nodal pressures. Moreover, at the next step, by deducing nodal pressure matrices of leakage scenarios and the primary matrix of nodal pressures (no consideration of leakage), and, at the final step at this section, differential pressure matrix of all nodes for all scenarios was determined. At the second section (steps 5 to 7), PPS were explored from following steps: pre-defined nodes (in this study, 10) that detects highest amount of pressure difference in each leakage scenario were found and selected. Then, these nodes were sorted based on their selection frequency among all leakage scenarios, and at the end, nodes that had the highest number of selection were considered as PPS (in this study, 105). At the third section (steps 8 to 11), PPSs were clustered into pre-defined categories (in this study, 3) using K-means clustering algorithm. This section started with implementing a sensor detection threshold of 0.7 psi in order to filter values of differential pressure matrix. Then, time-average values of differential pressure matrix (in this study, 97 hours[1]) for all leakage scenarios was calculated. After that, statistical values of time-average matrix including, maximum, minimum, average, and standard deviation for all leakage scenarios were computed. At the final step of this section, K-means algorithm is employed to cluster PPS into pre-defined categories based on previously calculated statistical values. Fig. S2 shows the K-means adaptation procedure for clustering potential pressure sensors.

[1] Simulation duration time of the hydraulic model was 97 hours

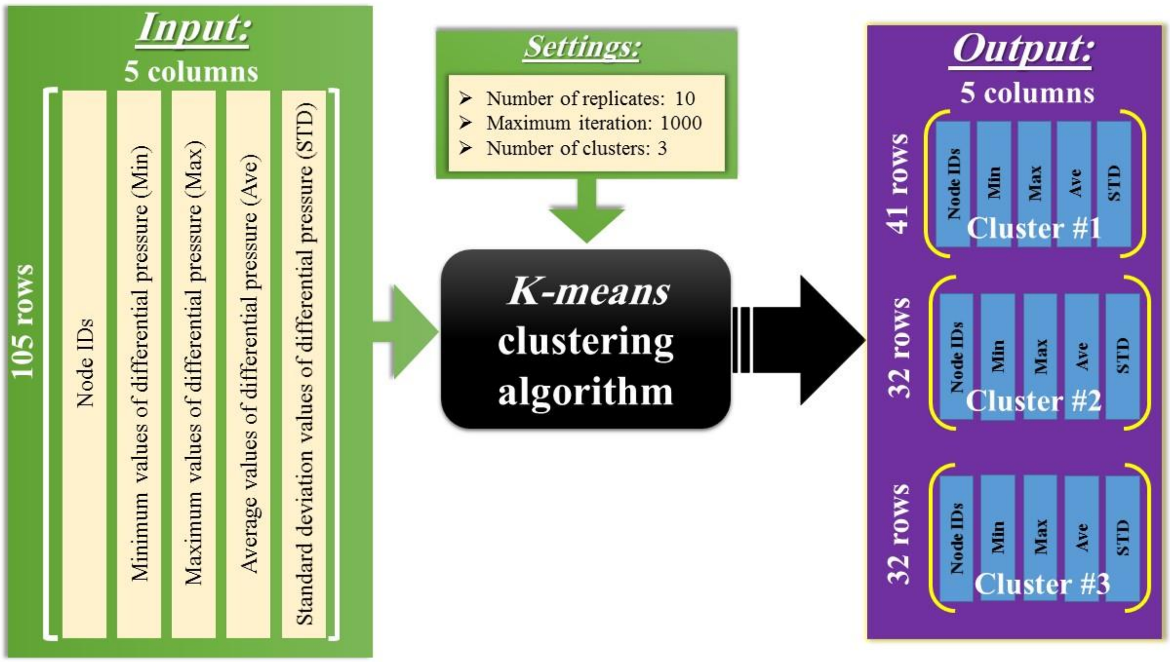


Fig. S2. K-means algorithm adoption procedure in proposed methodology in this study

At the fourth section (steps 12 and 13), transinformation entropy is developed to obtain Transinformation-Distance (T-D) curves for each potential pressure sensors' category. Subsequently, maximum and optimal distances (from Transinformation point of view) of each pressure sensors' category were determined based on developed T-D curves. At the fifth section, number of PPSs and their locations for each category were optimized using NSGA-II multi-objective optimization model. Main objectives of major stakeholders were defined as: 1. **M**aximizing the **C**overage of sensor network (MC), 2. **M**inimizing **E**xtra **I**nformation of sensor's **D**ata (MEID), and 3. **M**inimizing **N**umber of **S**ensors (MNS). Then, NSGA-II multi objective optimization was developed based on the proposed hybrid transinformation entropy-clustering model. Eventually, pareto optimal solutions for each category were acquired using the optimization model. At the final section, appropriate solution for number and location of pressure sensors in each category were found by implementing ELECTRE multi-criteria decision making model.

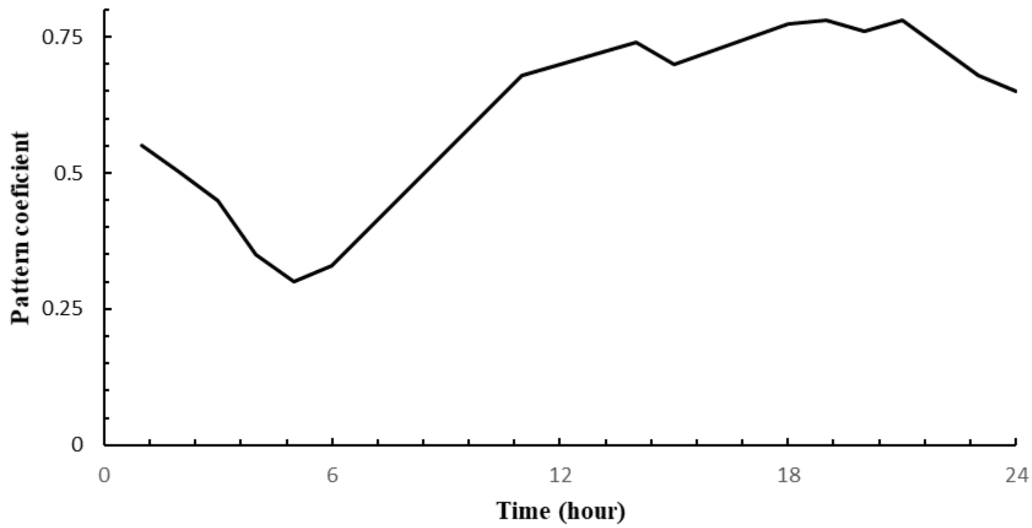


Fig. S3. Average values of pattern coefficient (diurnal pattern) of C-town WDS over commercial and residential demand patterns. (Adopted from Raei et al. 2018)

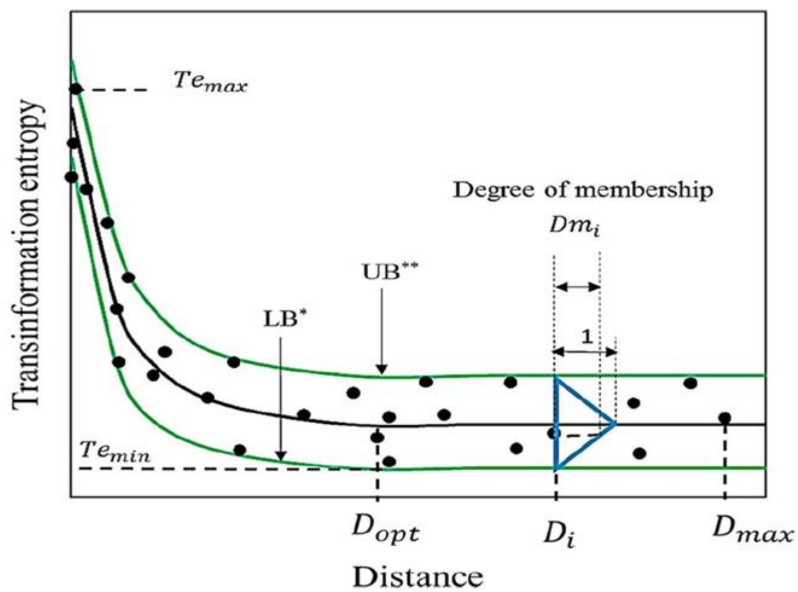


Fig. S4 Scheme of Traninformation and Distance (T-D) curve. (Adopted from Masoumi and Kerachian 2010)

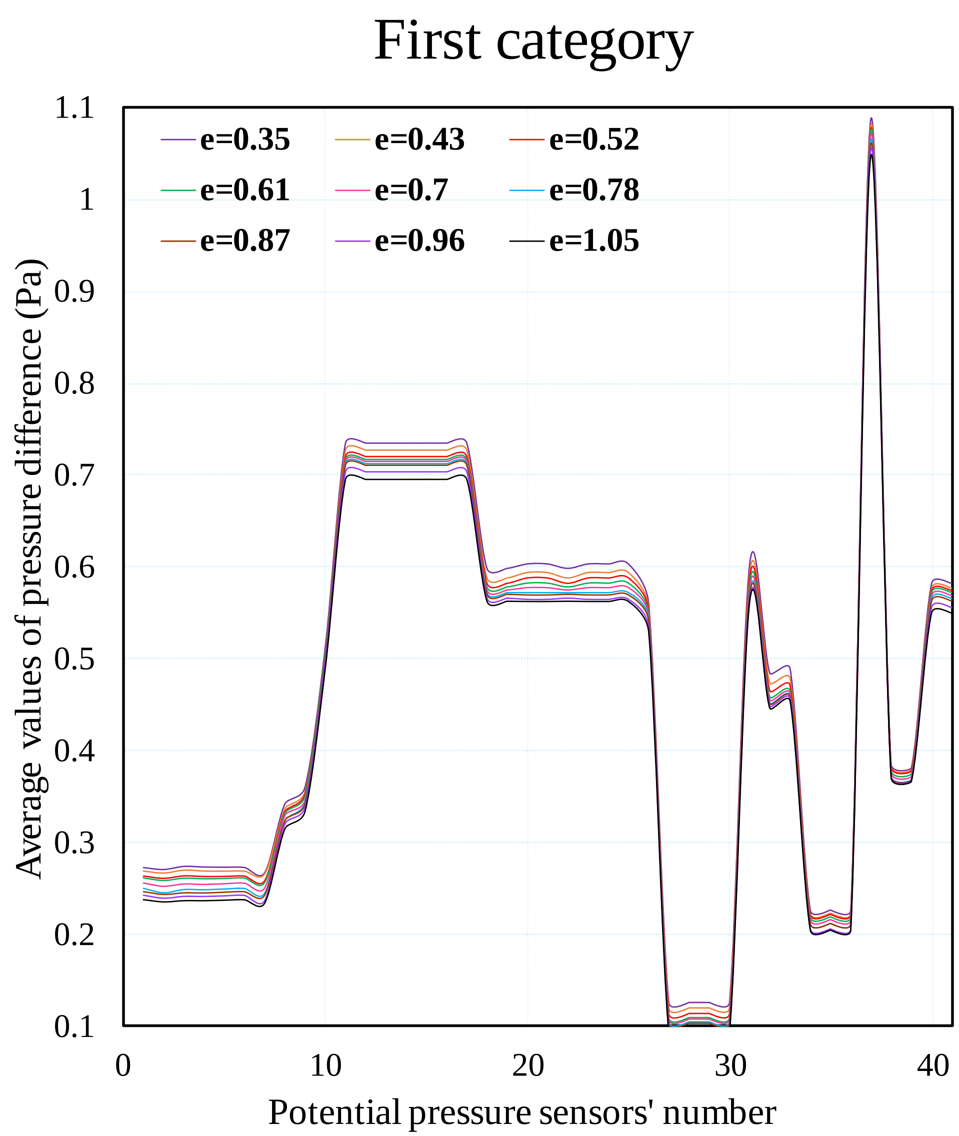
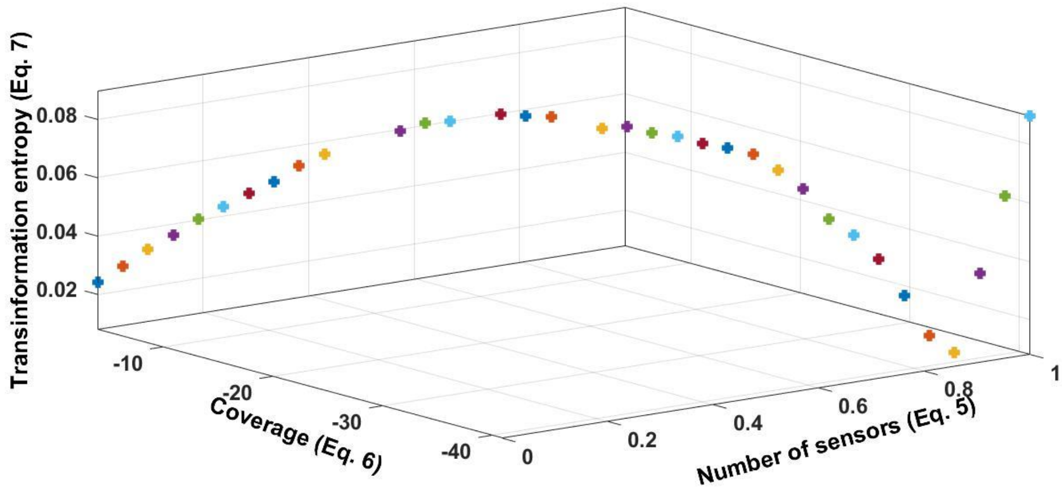
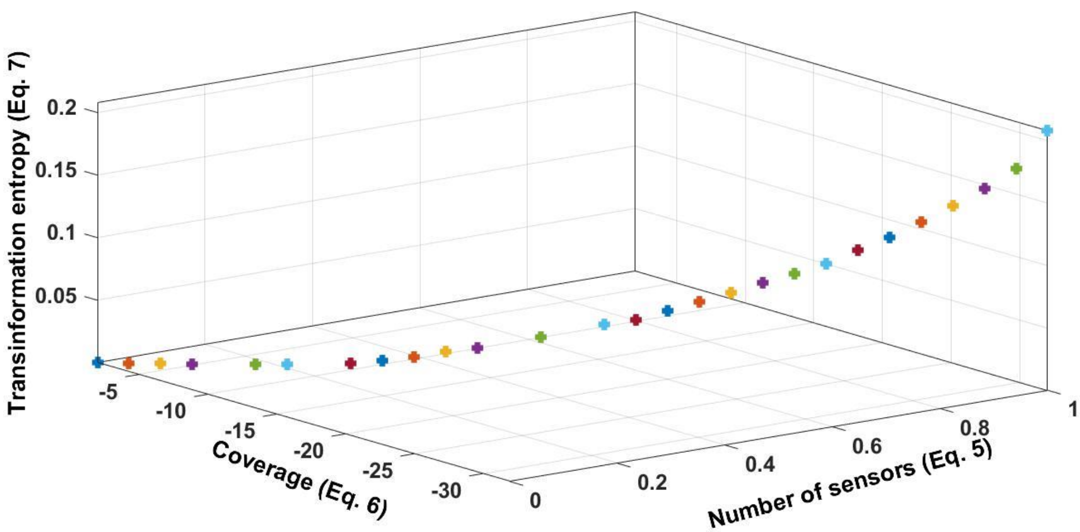


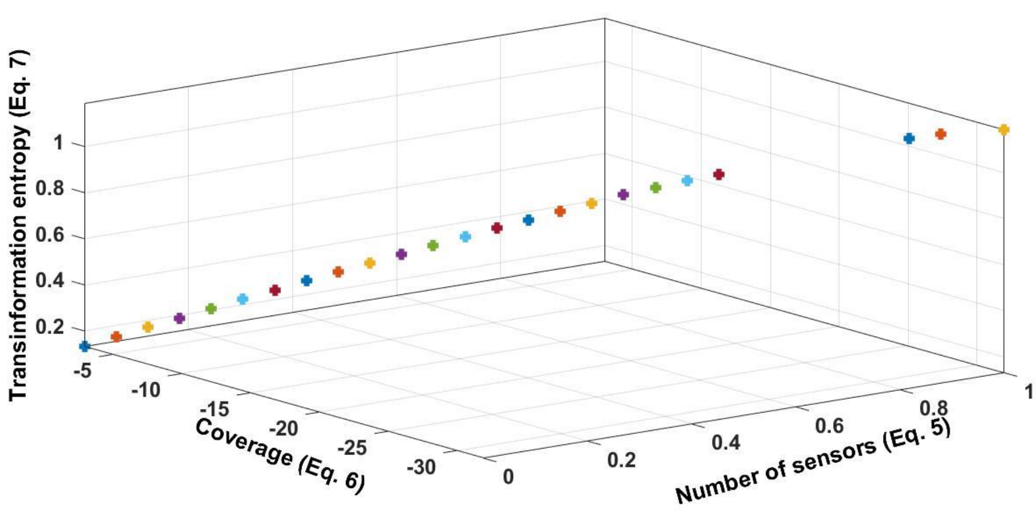
Fig. S5. Comparison of average values of pressure differences for the first category for all error thresholds of C-Town WDS


Transinformation entropy (Eq. 7)
0.08
0.06
0.04
0.02
-10
-20
-30
-40
Coverage (Eq. 6)
0
0.2
0.4
0.6
0.8
1
Number of sensors (Eq. 5)

(a)

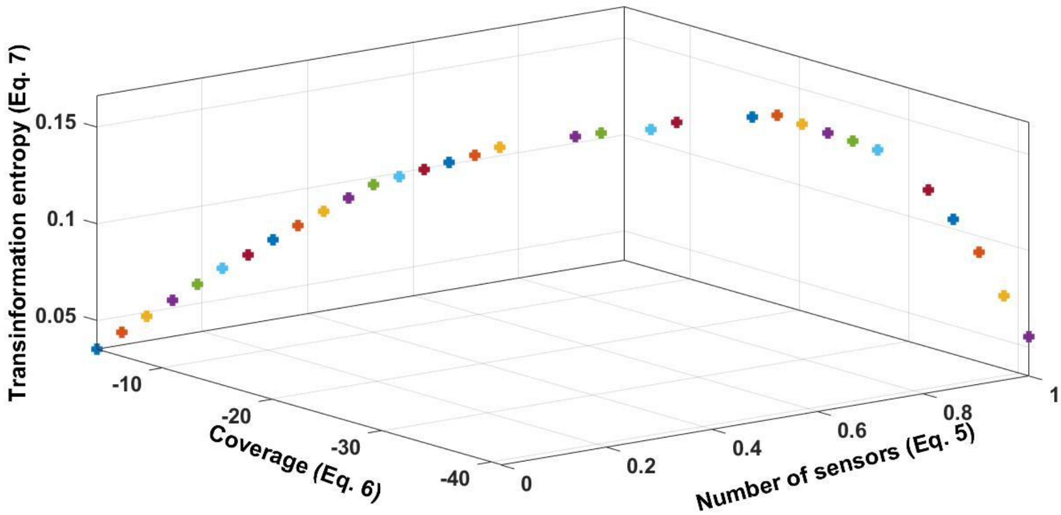

Transinformation entropy (Eq. 7)
0.2
0.15
0.1
0.05
-5
-10
-15
-20
-25
-30
Coverage (Eq. 6)
0
0.2
0.4
0.6
0.8
1
Number of sensors (Eq. 5)

(b)

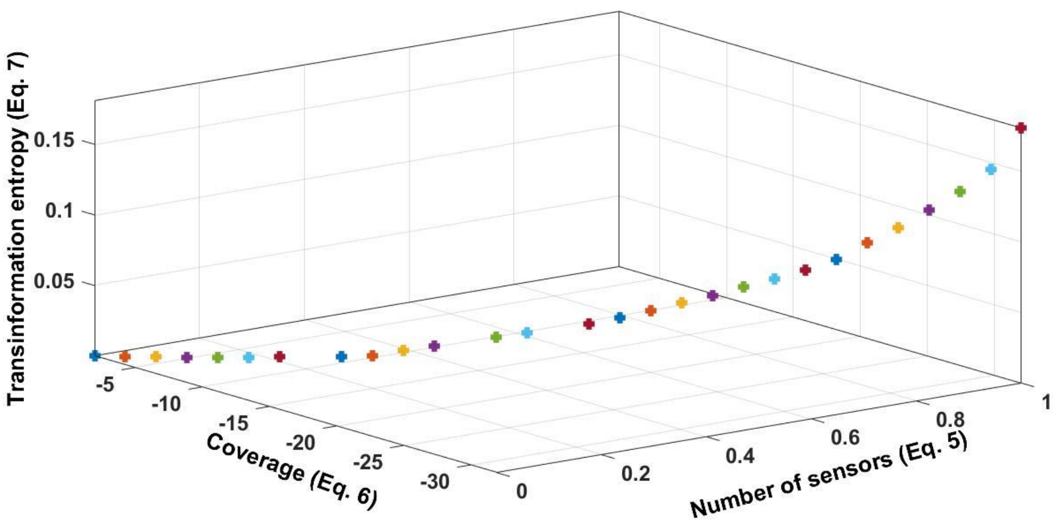

Transinformation entropy (Eq. 7)
1
0.8
0.6
0.4
0.2
-5
-10
-15
-20
-25
-30
Coverage (Eq. 6)
0
0.2
0.4
0.6
0.8
1
Number of sensors (Eq. 5)

(c)

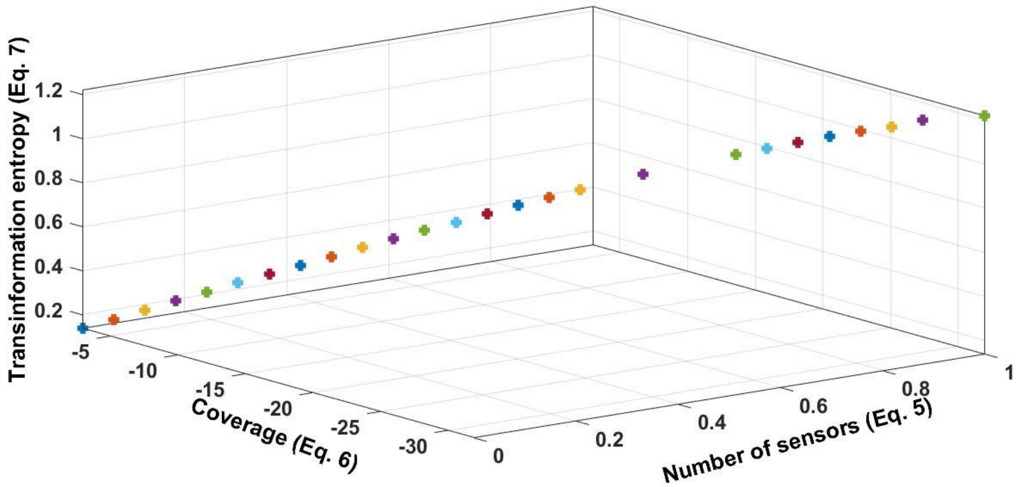
Fig. S6$_1$. Three-dimensional trade-off curve among objective functions categories for (a) first category, (b) second category and (c) third category corresponding to e=0.35



(a)

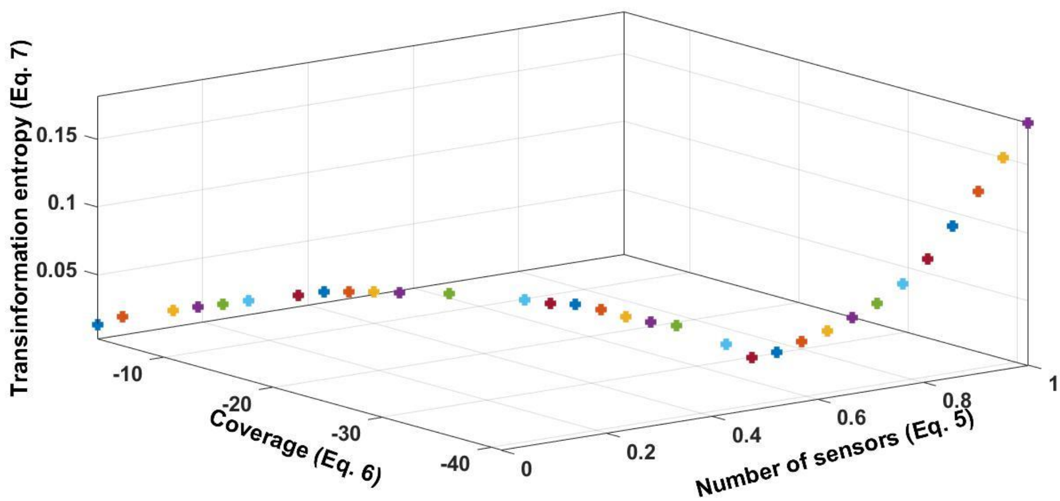


(b)

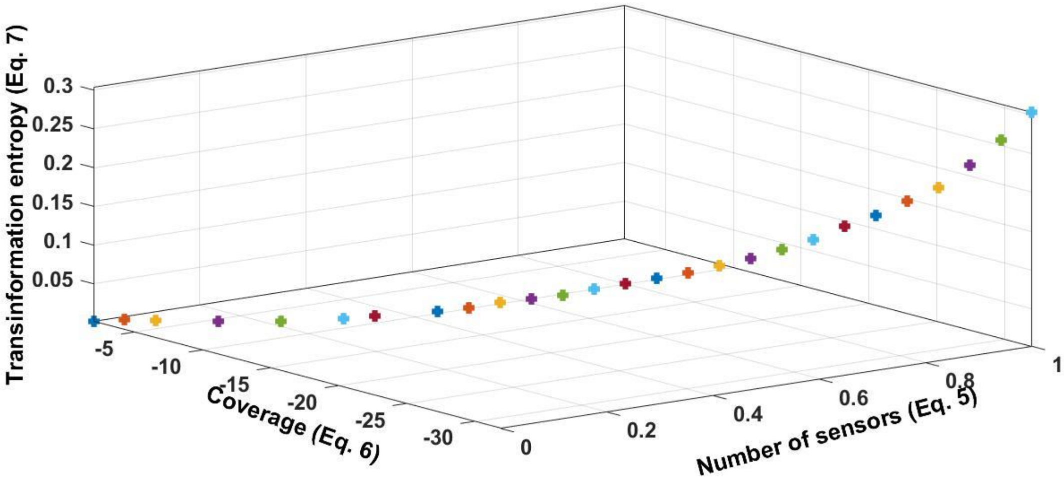


(c)

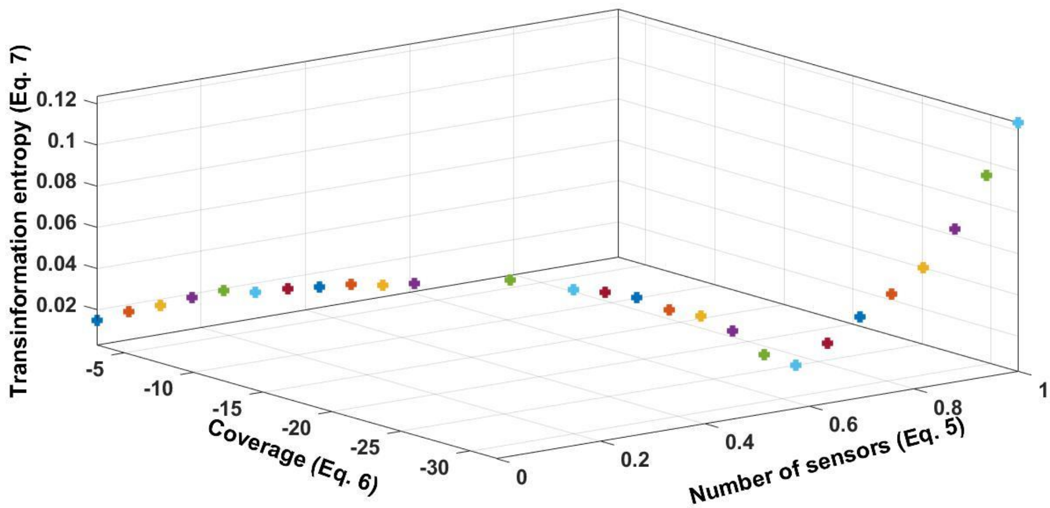
Fig. S6₂. Three-dimensional trade-off curve among objective functions categories for (a) first category, (b) second category and (c) third category corresponding to e=0.43



(a)

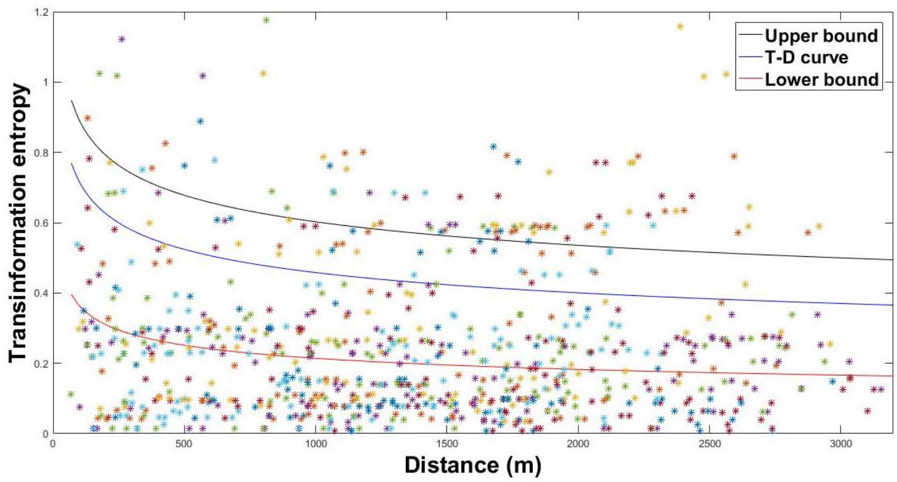


(b)

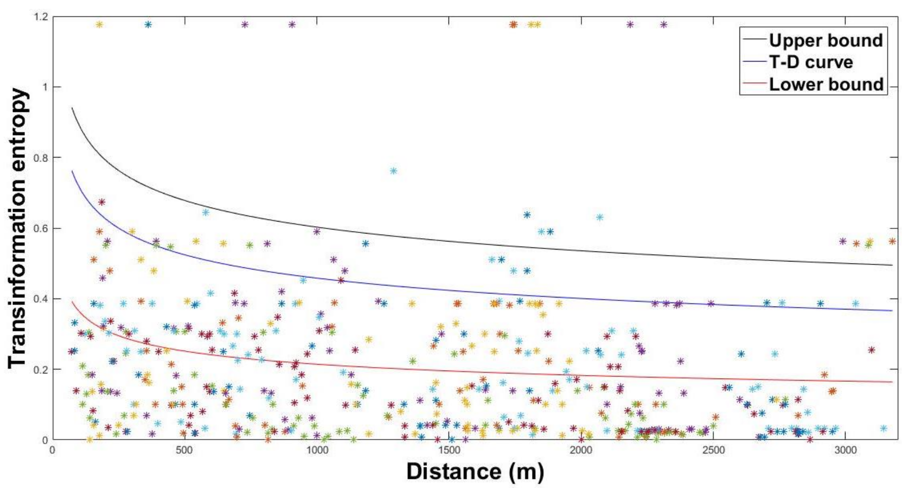


(c)

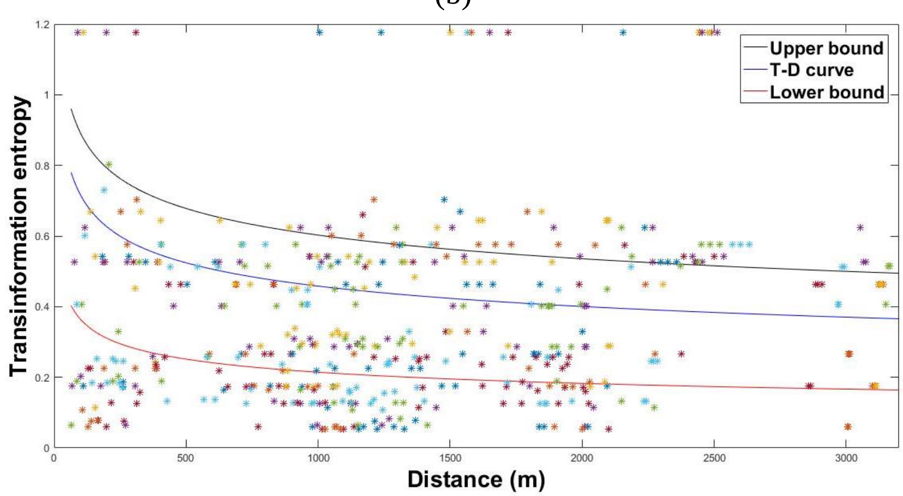
Fig. S6$_3$. Three-dimensional trade-off curve among objective functions categories for (a) first category, (b) second category and (c) third category corresponding to e=0.52


Upper bound
T-D curve
Lower bound
Transinformation entropy
Distance (m)
0
0.2
0.4
0.6
0.8
1
1.2
500
1000
1500
2000
2500
3000

(a)

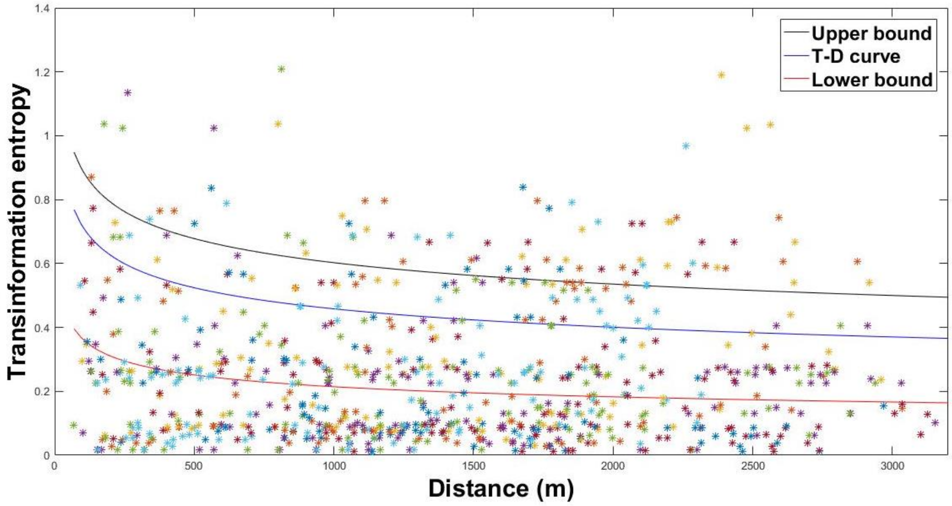

Upper bound
T-D curve
Lower bound
Transinformation entropy
Distance (m)
0
0.2
0.4
0.6
0.8
1
1.2
500
1000
1500
2000
2500
3000

(b)

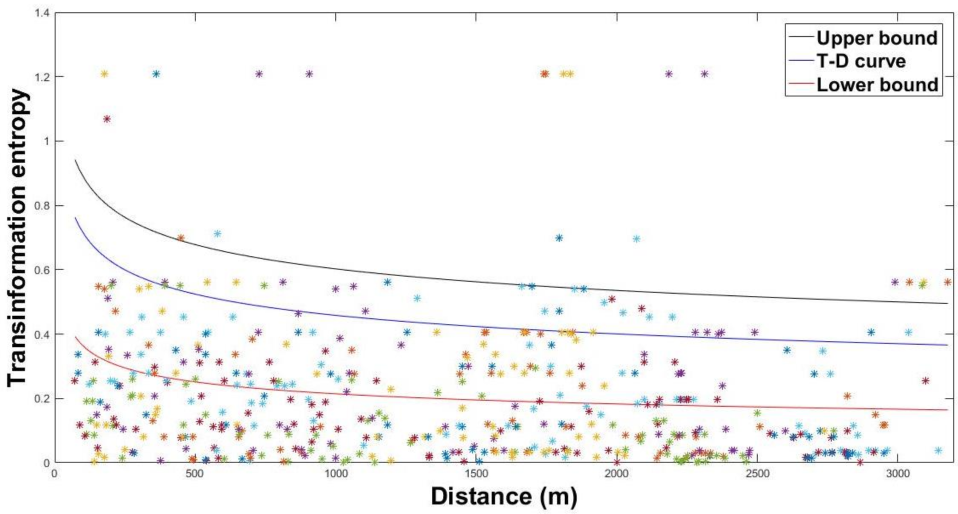

Upper bound
T-D curve
Lower bound
Transinformation entropy
Distance (m)
0
0.2
0.4
0.6
0.8
1
1.2
500
1000
1500
2000
2500
3000

(c)

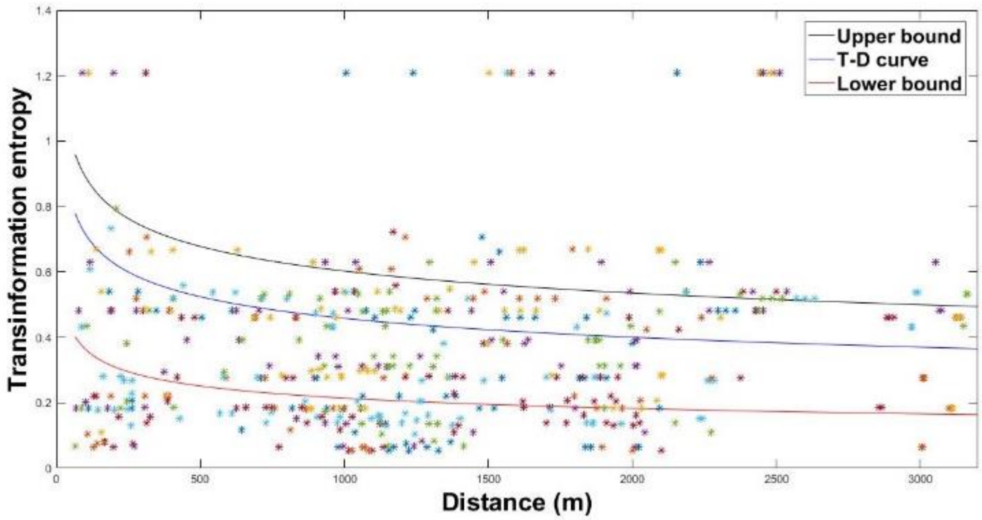
Figs. S6 Transinformation- Distance (T-D) curve for e=0.35 of C-Town WDS



(a)

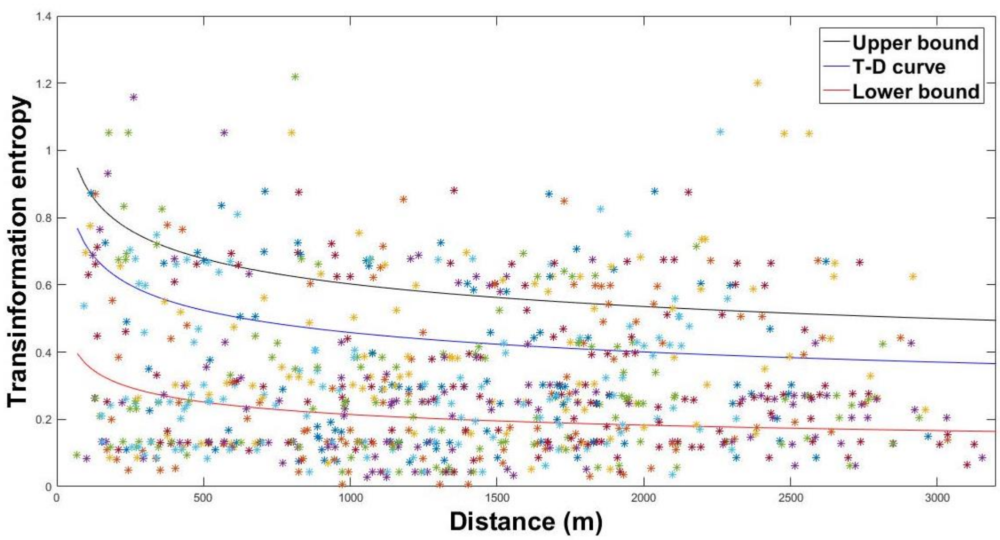


(b)

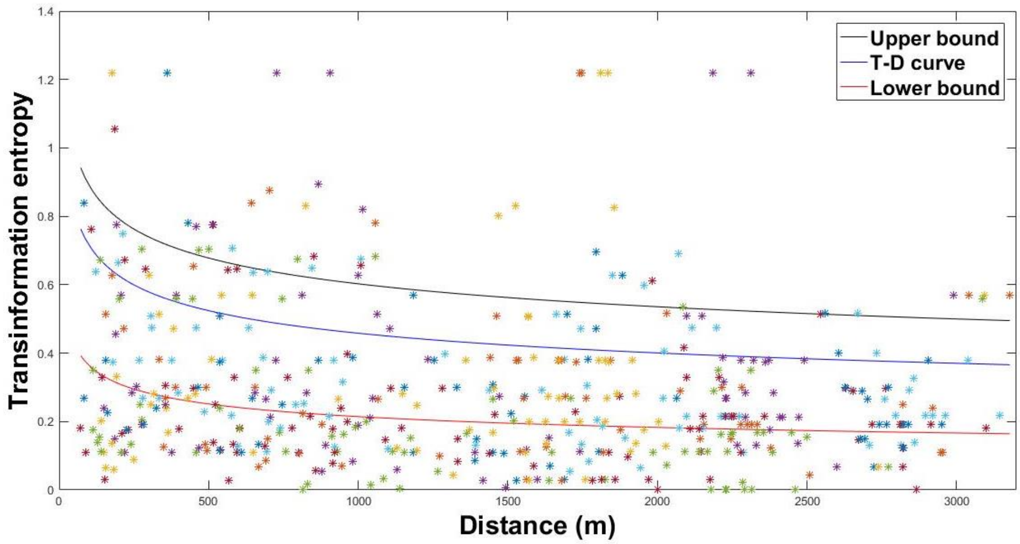


(c)

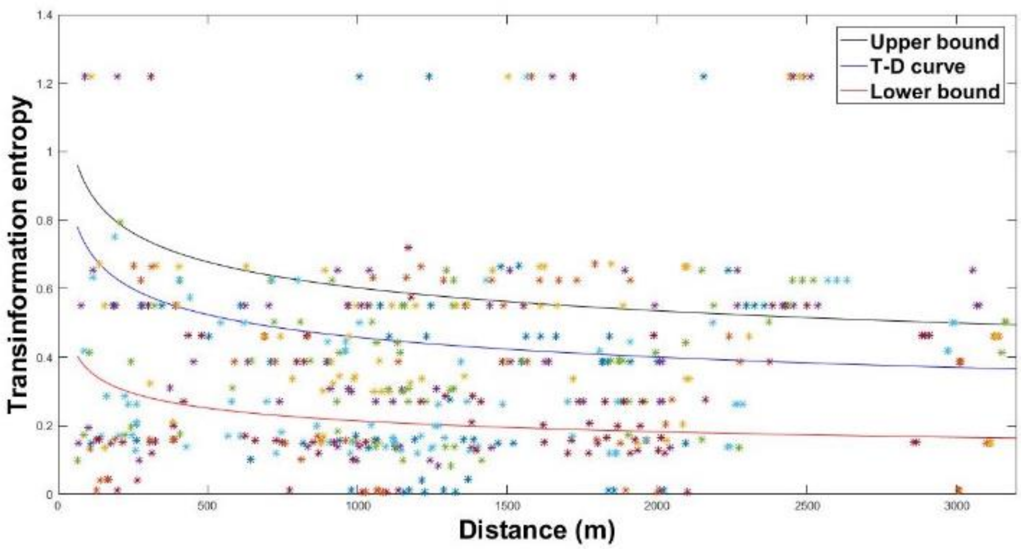
Figs. S6 Transinformation- Distance (T-D) curve for e=0.43 of C-Town WDS



(a)

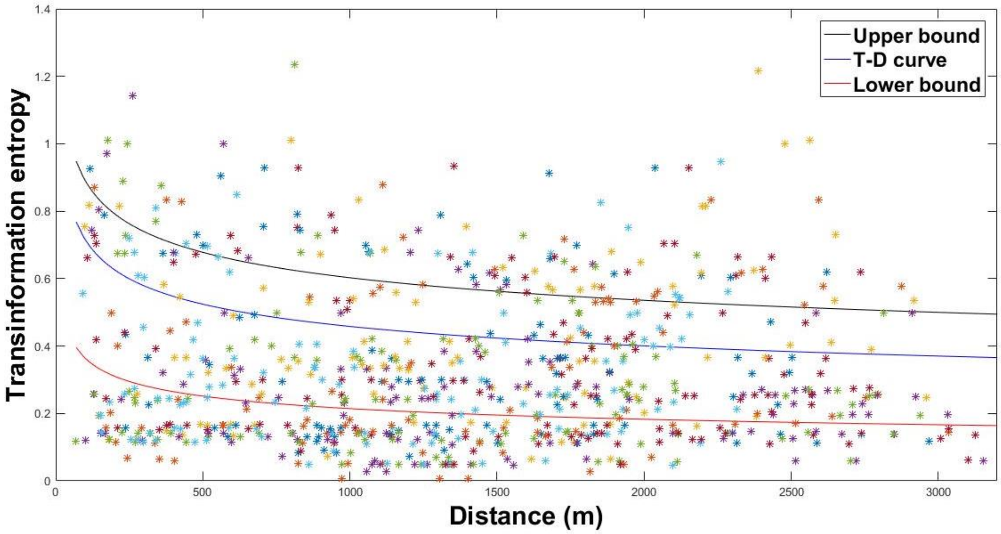


(b)

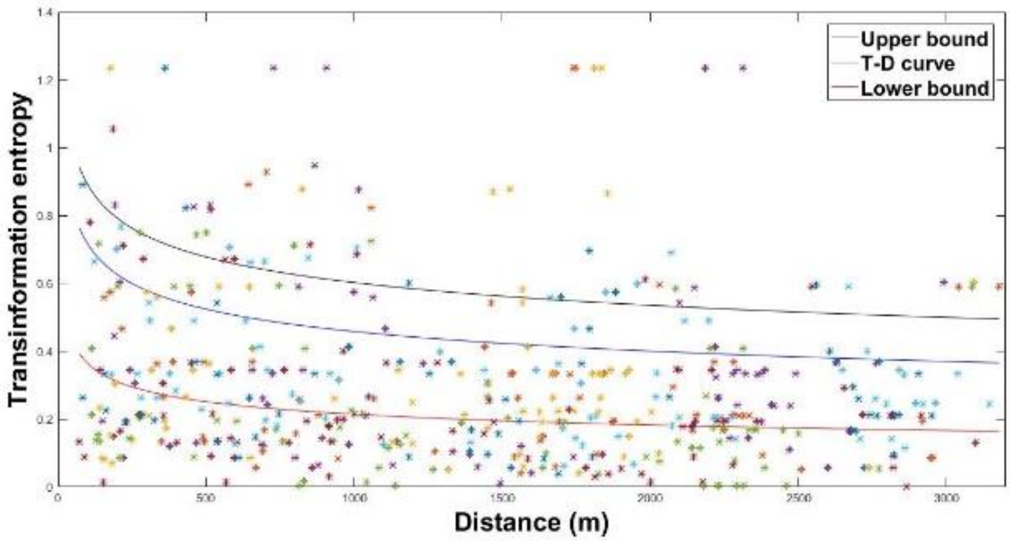


(c)

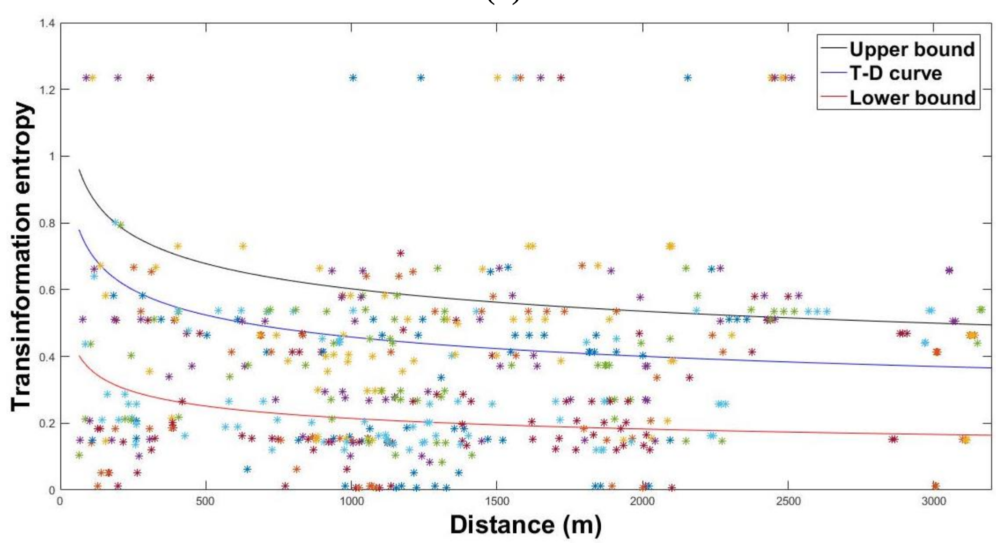
Figs. S6 Transinformation- Distance (T-D) curve for e=0.52 of C-Town WDS


Upper bound
T-D curve
Lower bound
Transinformation entropy
Distance (m)

(a)

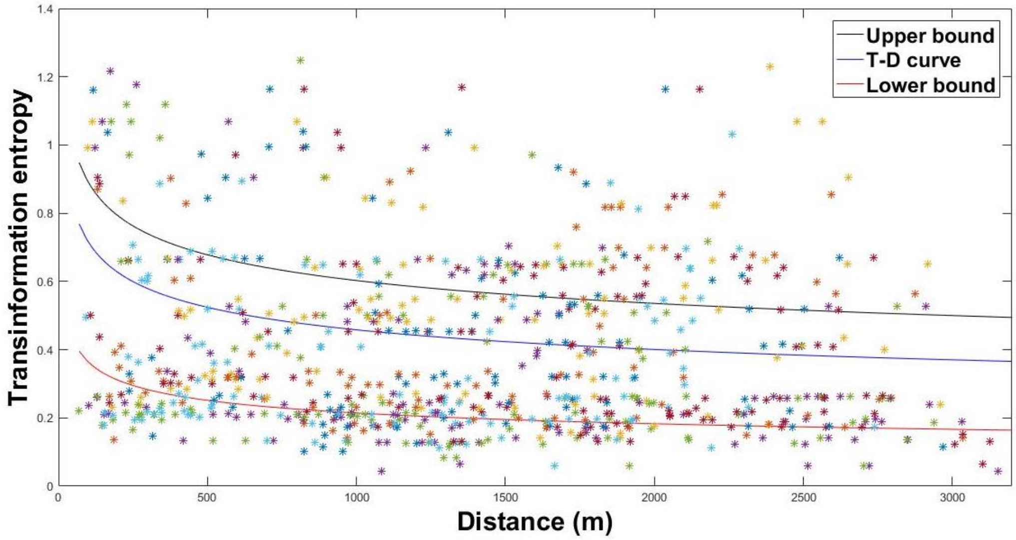

Upper bound
T-D curve
Lower bound
Transinformation entropy
Distance (m)

(b)

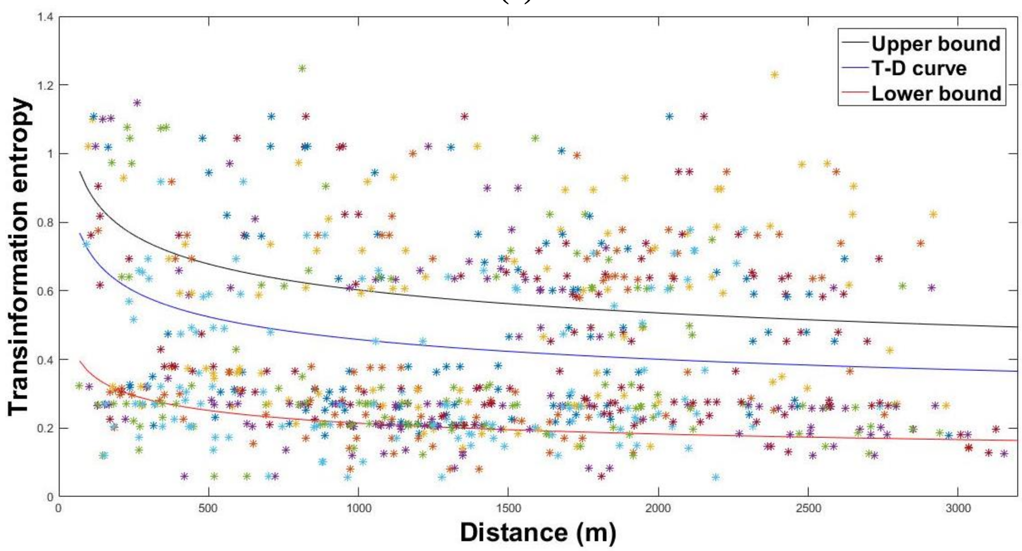

Upper bound
T-D curve
Lower bound
Transinformation entropy
Distance (m)

(c)

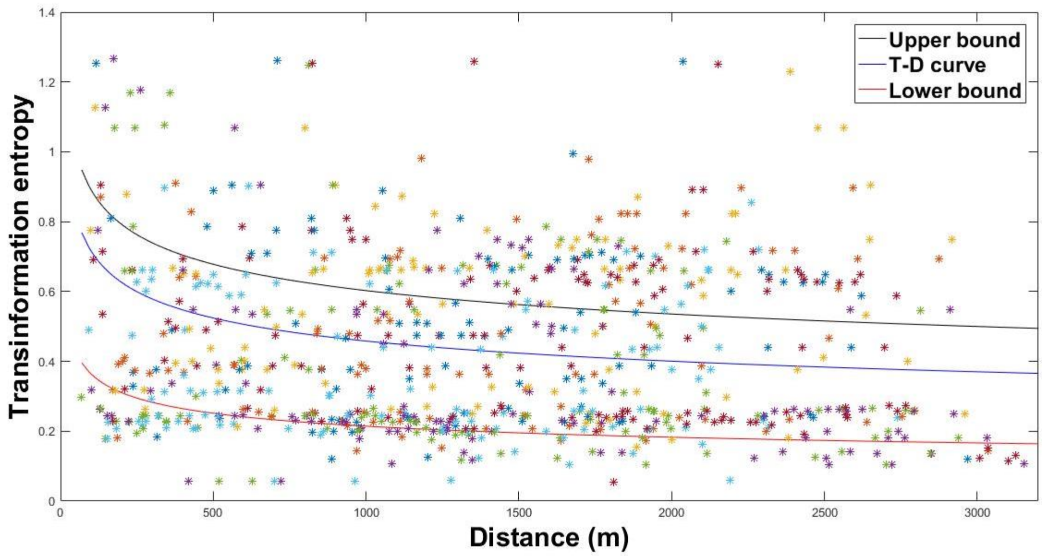
Figs. S6 Transinformation- Distance (T-D) curve for e=0.61 of C-Town WDS



(a)

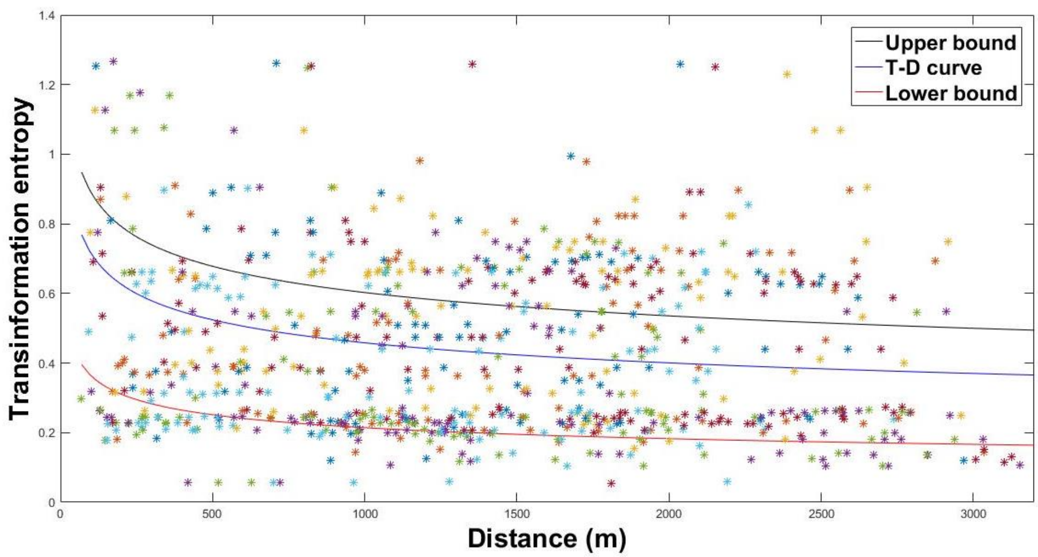


(b)

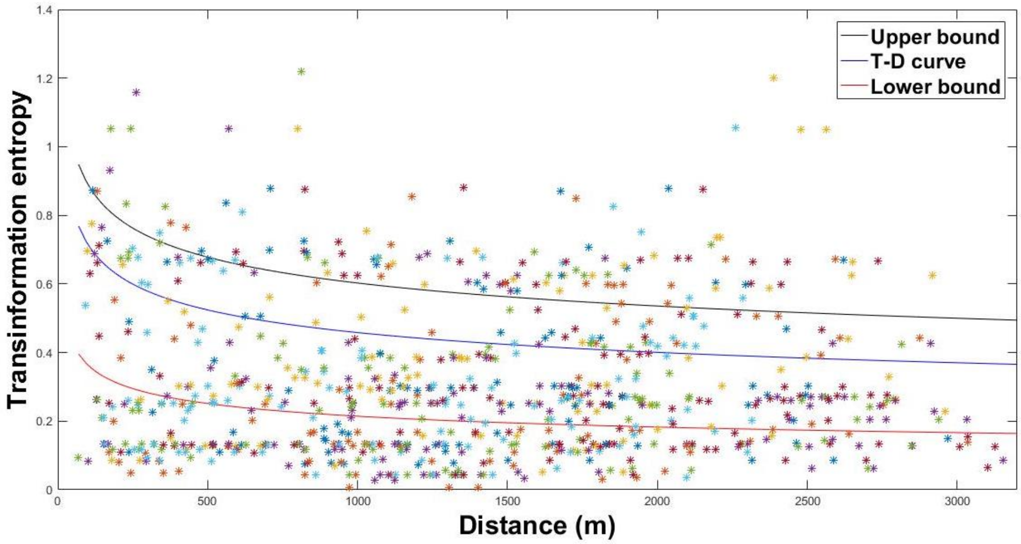


(c)

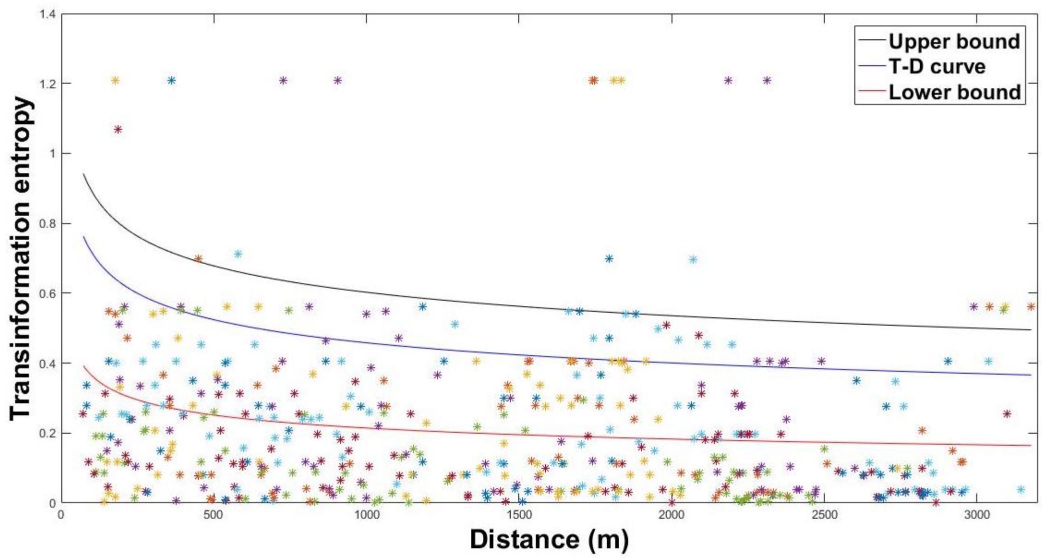
Figs. S6 Transinformation- Distance (T-D) curve for e=0.78 of C-Town WDS



(a)

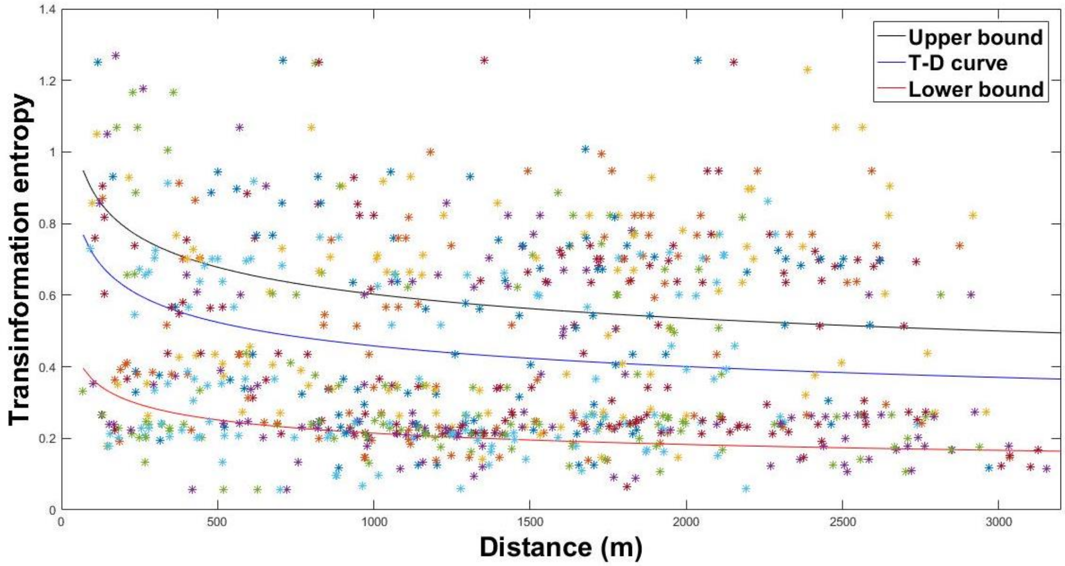


(b)

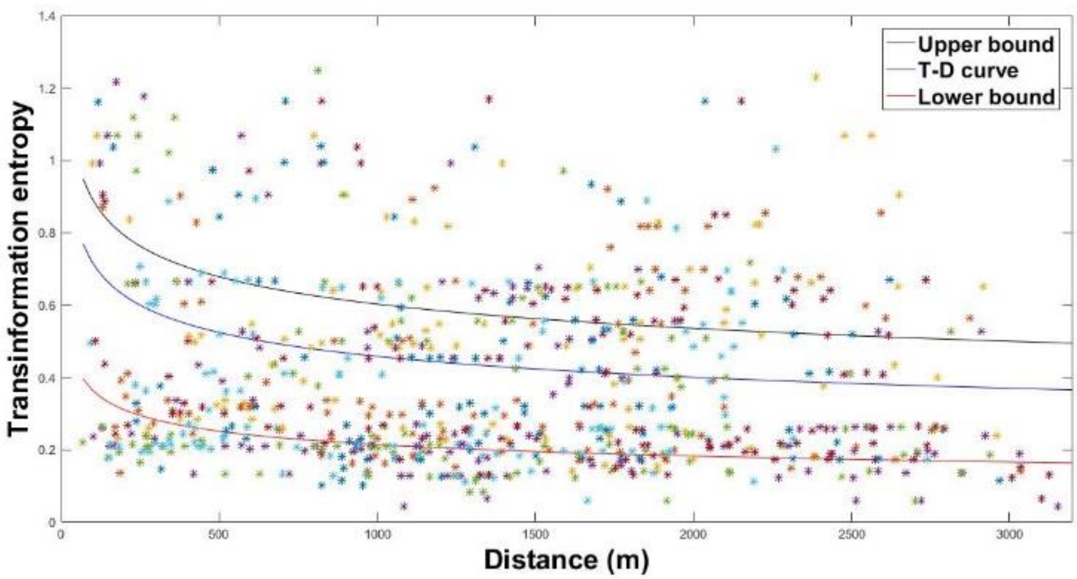


(c)

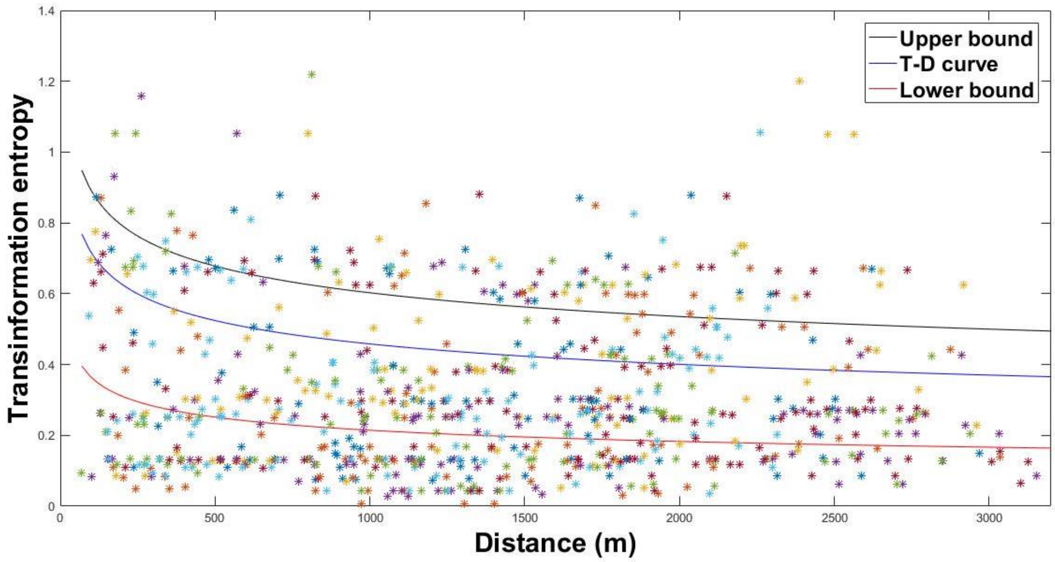
Figs. S6 Transinformation- Distance (T-D) curve for e=0.87 of C-Town WDS


Upper bound
T-D curve
Lower bound
Transinformation entropy
Distance (m)

(a)

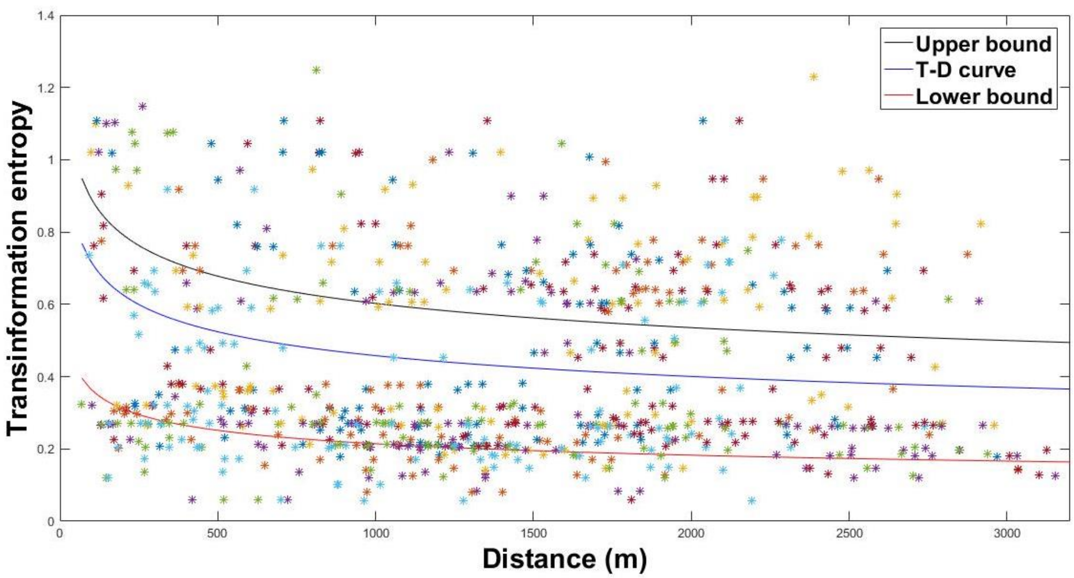

Upper bound
T-D curve
Lower bound
Transinformation entropy
Distance (m)

(b)

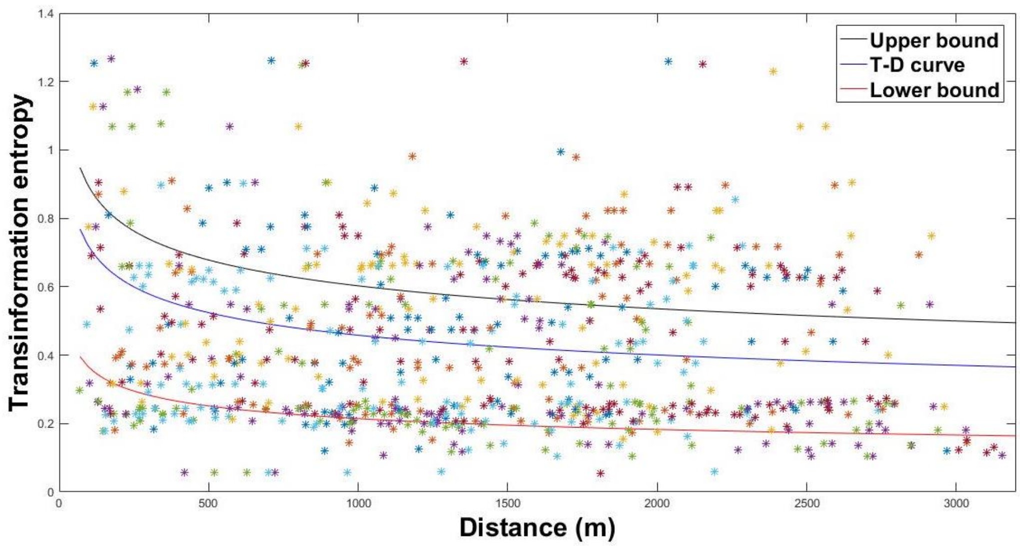


(c)

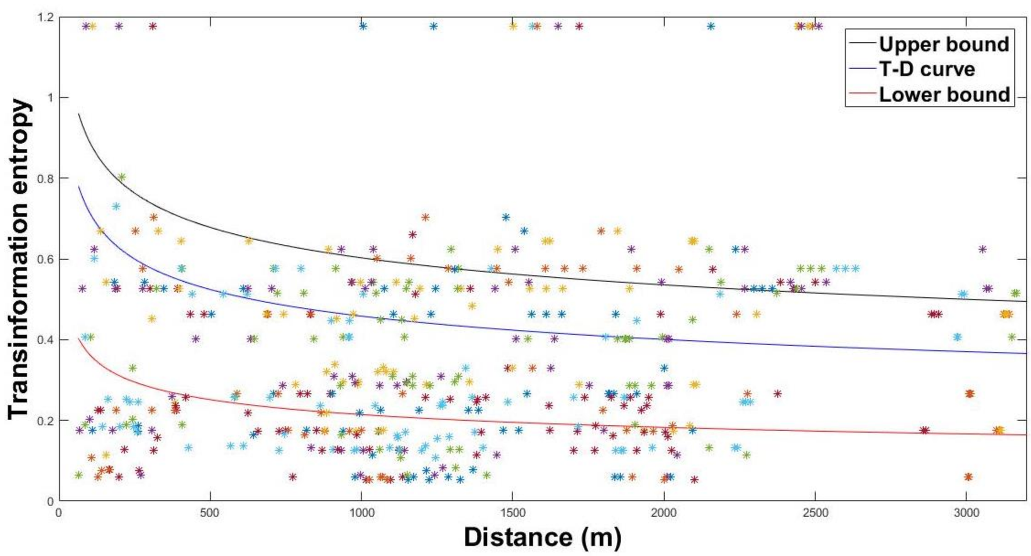
Figs. S6 Transinformation- Distance (T-D) curve for e=0.96 of C-Town WDS



(a)



(b)



(c)

Figs. S6 Transinformation- Distance (T-D) curve for e=1.05 of C-Town WDS

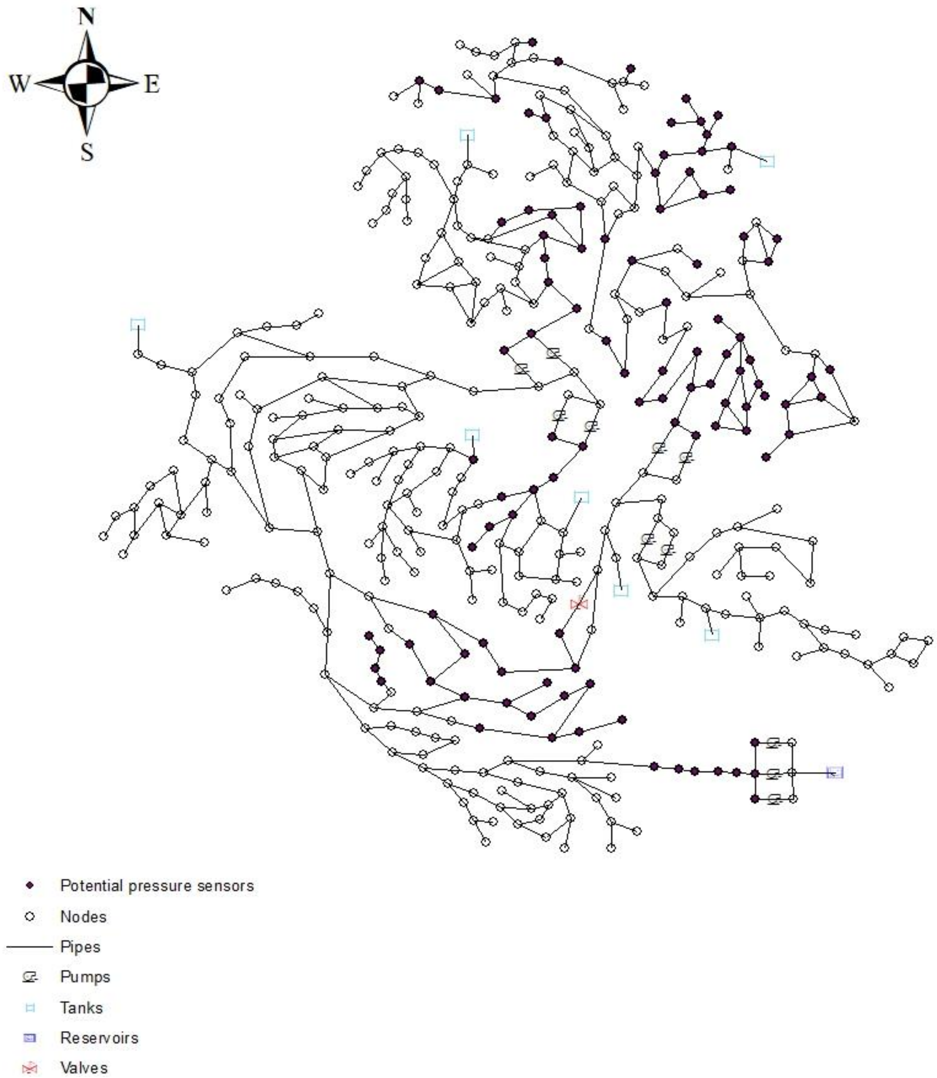


Fig. S7. Location of nodes and suggested potential pressure sensors in WDS of C-town

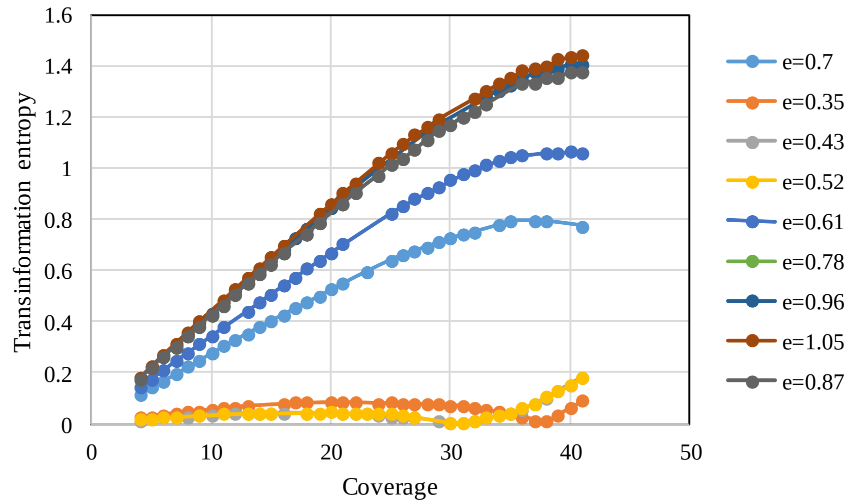


Fig. S8. Comparison of nine Pareto front of different error thresholds

Table S1. Maximum and optimal distances of each category of PPS based on T-D curves[1]

Distance (m)	Category #1	Category #2	Category #3
Maximum	3798	3177.6	3641.5
Optimum	1250	2100	1800

1. Transinformation- Distance

Table S2 Objective values of each proposed solutions by NSGA-II model

(a) First category for e=0.7

Solution number	Minimizing the Number of Sensors (Eq. 5)	Maximizing the coverage of sensor network (Eq. 6)	Minimizing extra information of pressure sensor's data (Eq.7)
1	0	4	0.110.14
2	0.02	4	0.16
3	0.05	5	0.16
4	0.055	6	0.19
5	0.081	6	0.22
6	0.108	7	0.24
7	0.135	8	0.27
8	0.612	9	0.30

9	0.162	10	0.32
10	0.189	10	0.35
11	0.21	11	0.37
12	0.24	12	0.40
13	0.27	13	0.42
14	0.29	14	0.45
15	0.32	15	0.47
16	0.35	16	0.50
17	0.37	17	0.52
18	0.0.4	18	0.54
19	0.43	19	0.59
20	0.45	20	0.64
21	0.51	21	0.66
22	0.56	23	0.67
23	0.59	25	0.69
24	0.59	26	0.71
25	0.62	27	0.73
26	0.64	28	0.75
27	0.67	29	0.77
28	0.70	30	0.79
29	0.72	31	0.79
30	0.75	32	0.79
31	0.81	34	0.79
32	0.83	35	0.77
33	0.89	37	0.77
34	0.91	38	0.77
35	1	41	0.76

(b) Second category

Solution number	Minimizing the Number of Sensors (Eq. 4)	Maximizing the coverage of sensor network (Eq. 5)	Minimizing extra information of pressure sensor's data (Eq.6)
1	0	2	0.006
2	0.03	4	0.01
3	0.066	5	0.01
4	0.10	6	0.013
5	0.133	7	0.017
6	0.1667	8	0.02
7	0.200	10	0.025
8	0.266	12	0.027
9	0.333	13	0.029
10	0.366	15	0.0270
11	0.433	16	0.016
12	0.466	17	0.22
13	0.500	19	0.024

14	0.566	19	0.016
15	0.600	20	0.010
16	0.566	22	0.006
17	0.700.	23	0.002
18	0.733	24	0.0024
19	0.766	24	0.0077
20	0.766	25	0.0.0154
21	0.800	26	0.0256
22	0.833	27	0.0384
23	0.866	28	0.0521
24	0.900	29	0.0686
25	0.933	30	0.0853
26	0.96	31	0.1104
27	1	32	0.1372

(a) Third category

Solution number	Minimizing the number of sensors (Eq.4)	Maximizing the coverage of pressure sensor network(Eq. 5)	Minimizing extra information of pressure sensors' data (Eq. 6)
1	0.0	2	0.14
2	0.033	3	0.21
3	0.066	7	0.28
4	0.1	8	0.35
5	0.133	9	0.42
6	0.167	10	0.49
7	0.167	11	0.56
8	0.20	11	0.63
9	0.23	12	0.70
10	0.26	13	0.75
11	0.30	14	0.770
12	0.33	15	0.84
13	0.40	16	0.91
14	0.43	19	0.98
15	0.46	20	1.04
16	0.60	21	1.011
17	0.63	23	1.11
18	0.70	24	1.37
19	0.760	25	1.43
20	0.79	26	1.55
21	0.80	27	1.60
22	0.840	28	1.66
23	0.86	30	0.73
24	0.96	31	1.77

25	1	32	1.83