Analyzing the historical rainfall observations and examining the flood event of Hopa, Turkey

SINAN JASIM HADI, MUSTAFA TOMBUL,
Department of Civil Engineering
Faculty of Engineering
Anadolu University
Eskisehir, Iki eylul Kampusu
Turkey
sinan.jasim@yahoo.com , mtombul@anadolu.edu.tr

Abstract: - Hopa a city of Artvin province in northeast of Turkey witnessed a devastating flood last year. The analysis of the event is conducted in this study by analyzing the rainfall time series data. Several distribution functions are fitted to the rainfall data and the best fitting function which identified by several tests used for constructing intensity - duration – frequency (IDF) curve. A formula representing the IDF curve developed using non-linear least square regression. The returning periods of the recorded rainfall intensities on the day of the event are obtained using the developed formula. Gumbel found as the best fitting function and used for IDF construction. The developed formula gives a tight agreement with a correlation 0.994 between the observed and predicted intensities. The highest return period calculated for the event observations using the formula is 211 years which belongs to the 4 hours' storm duration. The return period of 24 hours' duration which is the duration of the rainfall caused the flood is 41 years. The intensities of the flood event found smaller than the intensities of the observations recorded in 1988 for the durations ≤ 4 hours and higher for the durations > 4 hours.

Key-Words: - hopa, intensity-duration-frequency, rainfall, Gumbel, flood, trend analysis, stationarity.

1 Introduction

In the future, the intensity and the frequency of the extreme rainfall events are expected to witness a rise in the areas that already have frequent and intense events whereas areas with less frequent and intense are to witness a diminution [1]. Detecting the historical change not only in precipitation data but in any time series data can be done using trend and stationary analysis. Stationarity or non-stationarity is crucial for extreme rainfall events because it is the main assumption of the frequency analysis of the extreme rainfall events [2-4].

One of the statistical methods of extreme rainfall events analysis is Intensity – Duration – Frequency (IDF) curve. IDF defined by [5, 6] as a diagram illustrating the intensity of the rainfall falling on a basin for a specified period of time. IDF is used for extracting the rainfall intensity for various storm durations and several return periods. IDF is very important in the reduction of life and property losses as it useful in several aspects such as: assessing and judging the hazards, the damage occurred, and the preventive methods [7].

Based on the IDF curves, the mathematical relationship among rainfall intensity I, duration d, and return period T (also known as the frequency) can be developed [6, 8, 9]. This relationship can be used as an alternative for the IDF curve for the calculations of any of the missing variable. For example, in case of needing the intensity known as storm design this formula can be used that intensity can be obtained by substituting any return period and duration.

Many studies have been conducted around the world for constructing the IDF curves and developing formula representing these curves. The following are some instances: [6, 7, 10-14]. Several studies have also been conducted in a number of cities around Turkey such as: [15-20].

Hopa is a district of Artvin province located in the northeast of Turkey and on the eastern Turkey coast of the Black sea. A flood happened on 24/08/2015 in Hopa caused 8 deaths, 3 missing and 17 injured. The general objective of this study is to examine, and analyze the historical rainfall observations and compare them with the flood event observations. The detailed objectives are: a) Implementing stationarity and trend tests on the precipitation time series; b) constructing the IDF based on the best fitting distribution; c) developing a formula representing the IDF; d) comparing the flood observations and calculating their return periods.

2 Study area and Data

The aim of this study to analyze rainfall observations that caused a flood in Hopa which led to several deaths, and injuries in addition to a vast destruction in the properties as some of the houses were almost totally covered by the water. The heavy rain caused not only flood but landslides in several parts of the area. Only one station located at the place of the flood at the coordinates 41°24'23.55"N and 41°26'35.57"E was used as the closest station to that station has no hydrological effect to the flooded area due to the long distance from that area.

Data was collected for the chosen metrological station located at hopa from General Department of Meteorology - Ministry of Forest and Water Affairs of Turkey. Data from 1965 to 2015 was collected for the duration of 5, 10, 15, 30, 60, 120, 180, 240, 300, 360, 480, 720, 1080, 1440 min. The observations of 2015 which is the year the flood happened was not included in the IDF and formula calculations in order to be used for the comparison and calculating the return period.

3 Methodology

Detecting increase or decrease in any time series historical data is very important especially for identifying the climate change effect. In this study, two trend tests were implemented: Mann-Kendall test (MK) and Cox and Stuart test. MK and Cox and Stuart tests' null hypothesis (H_0) is that no trend is present and the alternative hypothesis (H_a) is trend is present. P value calculated for the two tests and compared with the significance levels are shown in table 1.

Trend test helps in detecting the increase or decrease in the historical data and the detected change does not provide information about nonstationarity that important in IDF constructing. Therefore, nonstationarity analysis conducted in this study using two tests: Kwiatkowski–Phillips–Schmidt–Shin (KPSS) test and Phillips–Perron (PP) test. KPSS test's null hypothesis (H_0) is that time series data are stationary and the alternative hypothesis (H_a) is data are nonstationary while PP test has the contrast hypotheses (i.e. H_0: Nonstationary, and H_a: Stationary). The tests statistics for both tests which calculated and compared with the critical values are listed in table 2.

Generally, IDF curves are constructed through several steps [21, 22]. Initially, the historical records are fitted to one of the distribution function and that is done for every duration. In this study, six distribution functions used: Weibull 3, Normal, Log-Normal 3, Log-Pearson 3, Gumbel, and Log-Logistic 3. The identification of the best fitting distribution was conducted using: Chi –Square, Kolmogorov-Smirnov, and Bayesian Information Criterion. The results are shown in appendix 1. The second step is to use the distribution functions to calculate the intensities for every duration and chosen return periods. In this study, only the best fitting distribution used for calculating the intensities using return periods: 2, 5, 10, 25, 50, 100, and 500. The calculated intensities with their frequency factors are tabled in appendix 2 and the developed IDF with the logarithms of the intensities to obtain better resolution is shown in fig.1.

An empirical IDF formula was also developed which is used for calculating the intensities as a dependent variable by substituting the storm duration and the return period which are considered as independent variables. Therefore, a power low relation [6, 7, 9, 21-23] was obtained:

$$I_T = \frac{a\,(T_r)^m}{(t_d + b)^e} \qquad (1)$$

Where I_T the intensity, T_r the return periods, t_d storm duration, and (a, m, b, and e) are the fitting parameters. In this study the b parameter was eliminated from the equation as it has no effect due to its very small value. Using the Non- Linear Least square regression, the parameters were obtained. After obtaining the formula the correlation between the observed and the predicted values was calculated in addition to plot them for having a visual evaluation. The obtained parameters, formula, and correlation are shown in table 3. The observed-predicted plot illustrated in fig.2.

The observations of 2015 that represent the extreme rainfall values which were all recorded on the day of the flood, were removed from the IDF and formula calculations for using them for the comparison. The observations which collected as the depth of the rainfall were converted to intensities. These intensities along with the storm duration were substituted in the obtained formula for calculating the expected return periods. The observations of 1988 are the highest records in the studied period. Therefore, they were chosen for the comparison. The observations of 2015 with their return periods and the observations of 1988 were listed in table 4.

4 Result and discussion

Trend analysis is conducted in this study using two tests Mann-Kendall test (MK) and Cox and Stuart test. The P values of these two tests for all storm durations are shown in table 1. All the values for the two test are higher than the three significance levels; 0.1, 0.05, 0.01. These results mean that the null hypothesis is failed to be rejected in all durations which leads to making a decision that the time series data has no trend.

Table 1. Test statistics values for each duration of two Trend analysis tests: Mann-Kendall (MK) and Cox and Stuart

Test	5 min	10 min	15 min	30 min	1 hr	2 hr	3 hr	4 hr	5 hr	6 hr	8 hr	12 hr	18 hr	24 hr
MK Test	0.17	0.14	0.13	0.23	0.38	0.32	0.27	0.16	0.29	0.39	0.17	0.19	0.51	0.94
Cox and Stuart	0.99	0.97	0.99	0.97	0.78	0.88	0.5	0.88	0.78	0.78	0.78	0.78	0.5	0.22
Confidence levels for MK and Cox and Stuart test are: 0.1, 0.05, 0.01														

The result of the nonstationarity two tests Kwiatkowski–Phillips–Schmidt–Shin (KPSS) test and Phillips–Perron (PP) test are shown in table 2. KPSS test is two sided test which means that as long as the test statistic value less than the critical value, the null hypothesis failed to be rejected. The test statistics of all storm durations are less than the critical values of all significance levels leading to the conclusion that data are stationary. On the other hand, PP test is left sided that if the test value less than the critical value the null hypothesis rejected taking into mind that the null hypothesis is that data are nonstationary. Test values of all durations are less than the critical values leading to the result that null hypothesis rejected, thus, the data are stationary.

Table 2. Test statistics values for each duration of two Nonstationarity analysis tests: Kwiatkowski–Phillips–Schmidt–Shin (KPSS) and Phillips–Perron (PP).

Test	5 min	10 min	15 min	30 min	1 hr	2 hr	3 hr	4 hr	5 hr	6 hr	8 hr	12 hr	18 hr	24 hr
KPSS Test	0.247	0.314	0.318	0.2	0.088	0.061	0.092	0.054	0.049	0.039	0.03	0.03	0.04	0.092
PP Test	-5.983	-6.978	-6.723	-6.999	-7.41	-7.746	-7.51	-8.134	-7.93	-7.916	-7.171	-7.107	-7.846	-7.061
Confidence Level	KPSS	0.1	0.05	0.01	PP Test	0.1	0.05	0.01						
Critical value		0.347	0.463	0.739		-3.176	-3.495	-4.138						

After having stationary data and without trend, six distribution functions used for fitting the data: Weibull 3, Normal, Log-Normal 3, Log-Pearson 3, Gumbel, and Log-Logistic 3. Gumbel found as the best fitting function. IDF curve plotted in fig.1 using the intensities calculated using Gumbel function. Based on the IDF curve, intensity decreases as the storm duration increases for any return period while it increases as the return period increases for any storm duration.

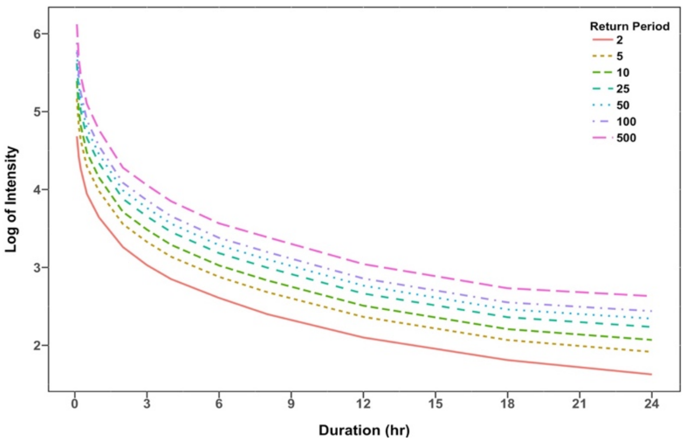

Fig.1 Fitted Intensity- Duration – Frequency (IDF) Curve of Gimbel function using extreme rainfall events for 14 durations storm and 6 return periods.

In order to ease the extraction of any of the variables included in the IDF curve; intensity, return period, and storm duration, a formula developed according to the relation between them. The parameters obtained for the formula shown in equation (1) are substituted in the formula to have a formula can calculate any of the missing variables having the other two. The result of developing this formula shown in table 3 reveals high correlation 0.994. A visual comparison between the observed that obtained from the fitting of Gumbel function listed in Appendix 1 and the predicted intensities shown in fig.2 illustrating tight agreement.

Table 3. Parameters of the developed formula and correlation.

Distribution	a	e	m	Derived Equation	Correlation
Gumbel	34.5	0.56	0.19	$I_T = \frac{34.5\,(T_r)^{0.19}}{(t_d)^{0.56}}$	0.994

Finally, the obtained formula used for calculating the return periods of the observation of 2015 which are recorded on the day of the flood event. The intensities converted from the collected depth and the calculated return periods for all storm durations shown in table 4. In the same table, the observations of 1988 the year that observed values are the highest among the years of the studied period are shown for the reason of comparison. The return period of the 24 hour is 41 years which means the intensity which is 12 mm/hr already recoded in the collected data. In 1988, the recorded intensity was 11 mm/hr which is close to that value. The highest return period is 211 years belongs to the duration 4 hours with an intensity 44.7 mm/hr and the value recorded in 1988 is 42.5 mm/hr.

Table 4. intensities of the year 2015 recorded on the day of the flood event, the intensities of 1988 representing the highest values in the studied period, and the return periods calculated by the developed formula.

Duration	5 min	10 min	15 min	30 min	1 hr	2 hr	3 hr	4 hr	5 hr	6 hr	8 hr	12 hr	18 hr	24 hr
Return period (Years)	1	2	2	4	11	48	106	211	153	112	67	65	48	41
Intensity (mm/hr)(2015)	120	105	82.8	68.0	55.1	49.6	46.0	44.7	37.1	31.6	24.3	19.3	14.5	12.0
Intensity (mm/hr)(1988)	606	363.6	282.8	181.8	125	67.5	53.8	42.5	34.0	28.4	21.3	14.2	11.6	11.0

In general, the intensities recorded in 1988 are higher than those recorded in 2015 in the duration ≤ 4 hours, while for the durations > 4 hours the 2015 intensities are higher.

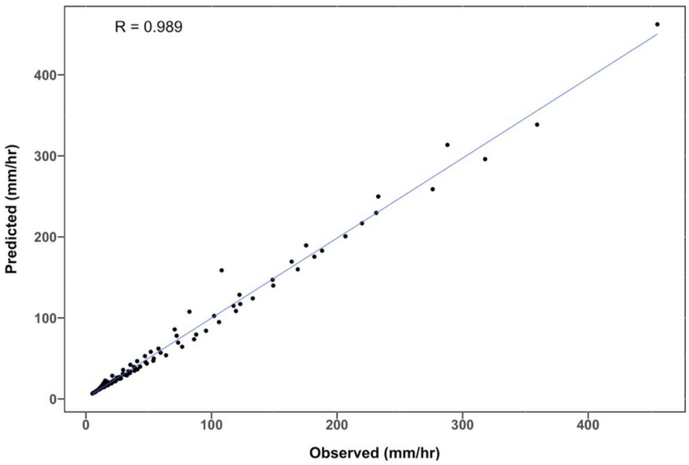

Fig.2 The observed values (i.e. obtained from the fitting of Gumbel function listed in Appendix 1.) vs the predicted values obtained from the developed formula.

5 Conclusion

The main aim of this study is to analyse the historical rainfall time series data of Hopa and the observations recorded on the day of the flood and compare them to have a broad idea about the event. The rainfall data proved as a stationary data and have no significant trend for the studied period 1965 – 2014. The observations recorded on the event day 23-24/08/2015 found as very close to the observations recorded in 1988 especially in the storm duration > 4 hours while for the durations ≤ 4 hours the event records are smaller. Although, the return period of 4 hours is found 211 years but the intensity recorded in the event 44.7 mm/hr is very close to the intensity 42.5 mm/hr which is recorded in 1988. The return periods 24 and 18 hours' duration are 41 and 48 years respectively which considered as not high values.

The calculation of the return periods and the comparison with the observations recorded in the studied period proved that the recorded intensities are expected to return with in the near future. the intensities recorded 28 years ago did not cause a flood and landslides like the one happened on the events day.

This study has not shown the reason of the flood. So, there is a shortcoming in this study that the analyses implemented based on one station records and that could be not enough to have a full image of the event. Therefore, there are two highly recommended points: spatial analysis of the event covering the entire area that the rainfall fell on, hydrologic modelling of the rainfall and flooded area with including the topography and the existing infrastructure.

Acknowledgments: - The authors would like to thank (DMİ) Devlet Meteroloji İşleri (General Department of Meteorology - Ministry of Forest and Water Affairs- Turkey) for providing the data to complete this study.

References

[1] IPCC, *Climate Change 2007: The Physical Science Basis. Contribution ofWorking Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge and NewYork: Cambridge University Press, 2007.

[2] a. G. Yilmaz and B. J. C. Perera, "Extreme Rainfall Non-Stationarity Investigation and Intensity-Frequency-Duration Relationship," *Journal of Hydrologic Engineering,* 2013.

[3] A. G. Yilmaz, M. A. Imteaz, and B. J. C. Perera, "Investigation of non-stationarity of extreme rainfalls and spatial variability of rainfall intensity–frequency–duration relationships: a case study of Victoria, Australia," *International Journal of Climatology,* 2016,

[4] E. A. Rosenberg, P. W. Keys, D. B. Booth, D. Hartley, J. Burkey, A. C. Steinemann*, et al.*, "Precipitation extremes and the impacts of climate change on stormwater

infrastructure in Washington State," *Climatic Change,* vol. 102, 2010, pp. 319-349.

[5] B. S. Dupont and D. L. Allen, "Revision of the rainfall intensity duration curves for the commonwealth of kentucky," USA2000.

[6] I. H. Elsebaie, "Developing rainfall intensity–duration–frequency relationship for two regions in Saudi Arabia," vol. 24, ed, 2012, pp. 131-140.

[7] S. A. AlHassoun, "Developing an empirical formulae to estimate rainfall intensity in Riyadh region," vol. 23, ed, 2011, pp. 81-88.

[8] D. Koutsoyiannis, "On the appropriateness of the gumbel distribution in modelling extreme rainfall," University of Bologna, Bologna, 2003
, pp. 303-319.

[9] D. Koutsoyiannis, D. Kozonis, and A. Manetas, "A mathematical framework for studying rainfall intensity-duration-frequency relationships," *Journal of Hydrology,* vol. 206, 1998, pp. 118-135.

[10] V. R. Baghirathan and E. M. Shaw, "Rainfall depth-duration-frequency studies for Sri Lanka," *Journal of Hydrology,* vol. 37, 1978, pp. 223-239.

[11] A. Gert, D. J. Wall, E. L. White, and C. N. Dunn, "Regional Rainfall intensity-duration-function curves for Pennsylvania," *Water Resources Bull,* vol. 23, 1987, pp. 479-486.

[12] T. A. Buishand, *Rainfall Depth-Duration-frequency curves a problem of dependent extremes*. Chichester: WILEY, 1993.

[13] T. A. Endreny and N. Imbeah, "Generating robust rainfall intensity-duration-frequency estimates with short-record satellite data," *Journal of Hydrology,* vol. 371, 2009, pp. 182-191.

[14] D. M. Hersfield, "Estimating the probable maximum precipitation," *Journal of the hydraulics Division,* vol. 87, 1961, pp. 99-116.

[15] R. Acar, S. Celik, and S. Senocak, "Rainfall Intensity-Duration - Frequency Model using an artificial neural network approach," *Journal of Scientific & industrial Research,* vol. 67, 2008, pp. 198-202.

[16] R. Acar and S. Senocak, "Modelling of Short Duration Rainfall (SDR) Intensity Equations for Ankara, Turkey," Republic of Macedonia, 2008, pp. 1-9.

[17] O. L. Asikoglu and B. Ertugrul, "Simple generalization approach for intensity–duration– frequency relationships," *HYDROLOGICAL PROCESSES,* vol. 28, 2014, pp. 1114-1123.

[18] T. Haktanir, M. Cobaner, and O. Kisi, "Frequency analyses of annual extreme rainfall series from 5 min to 24 h," *Hydrological Processes,* vol. 24, 2010, pp. 3574-3588.

[19] S. Senocak and R. Acar, "Modelling of Short Duration Rainfall (SDR) Intensity Equations for Erzurum, Turkey," *Journal of Engineering Science,* vol. 13, 2007, pp. 75-80.

[20] S. Senocak and R. Acar, "Modelling of short-duration rainfall intensity equations for the Agean region of Turkey," *FRESENIUS ENVIRONMENTAL BULLETIN,* vol. 16, 2007,

[21] W. T. Chow, *Handbook of Applied Hydrology*: McGraw-Hill, 1988.

[22] V. P. Singh, *Elementary hydrology*. New Jersey: Prentice Hall, 1992.

[23] L. M. Nhat, Y. Tachikawa, and K. Takara, "Establishment of intensity-duration-frequency curves for precipitation in the monsoon area of Vietnam," *Annuals of Dis. Prev. Res. Inst,* 2006, pp. 93-103.

Appendix 1. Calculation of frequency intensity I_T (mm/hr) values for different durations t_d (minutes and hours) and return periods *Tr* (years) using Gumbel method.

Duration	**5 Minute**			**10 Minute**			**15 Minute**			**30 Minute**			**1 Hour**		
Distribution	**CHS**	**KS**	**BIC**	**CHS**	**KS**	**BIC**	**CHS**	**KS**	**BIC**	**CHS**	**KS**	**BIC**	**CHS**	**KS**	**BIC**
Weibull3	0.000	0.266	275.5	0.003	0.146	317.3	0.002	0.150	344.2	0.055	0.118	384.5	0.550	0.107	421.7
Normal	0.000	0.272	333.5	0.000	0.211	350.4	0.000	0.190	367.2	0.017	0.168	399.7	0.303	0.108	434.2
Log-normal3	0.001	0.141	270.0	0.064	0.103	308.0	0.080	0.135	334.9	0.329	0.090	378.4	0.870	0.073	418.3
Log-Pearson 3	0.002	0.142	268.6	0.073	0.098	306.9	0.089	0.139	334.0	0.339	0.095	378.0	0.855	0.075	418.2
Gumbel	**0.022**	**0.118**	273.6	**0.218**	0.097	306.5	0.215	0.116	331.4	**0.549**	0.086	**374.7**	**0.964**	0.059	**414.6**
Log-Logis3	0.014	0.122	**264.3**	0.193	**0.075**	**302.9**	**0.345**	**0.114**	**328.6**	0.537	**0.072**	375.2	0.906	**0.052**	417.0
	2 Hour			**3 Hour**			**4 Hour**			**5 Hour**			**6 Hour**		
Weibull3	0.168	0.085	446.7	0.023	0.123	464.4	0.004	0.146	473.7	0.002	0.144	482.7	0.043	0.119	486.7
Normal	0.103	0.107	449.5	0.013	0.161	468.1	0.003	0.183	474.7	0.004	0.176	481.5	0.071	0.137	484.2
Log-normal3	0.348	0.058	444.2	0.156	0.096	458.7	0.084	0.118	467.7	0.042	0.127	477.3	0.243	0.100	482.7
Log-Pearson 3	0.366	0.073	445.4	0.191	0.083	459.0	0.105	0.107	468.7	0.044	0.104	480.4	0.257	0.112	487.2
Gumbel	**0.501**	0.063	**440.2**	**0.307**	0.089	**453.9**	0.185	0.115	**462.6**	0.095	0.115	**473.1**	0.402	0.100	**479.4**
Log-Logis3	0.367	**0.052**	443.5	0.286	**0.065**	456.3	**0.209**	**0.088**	464.2	**0.203**	**0.097**	474.0	**0.528**	**0.082**	480.6
	8 Hour			**12 Hour**			**18 Hour**			**24 Hour**					
Weibull3	0.004	0.105	498.6	0.017	0.131	503.8	0.002	0.120	509.5	0.188	0.101	510.6			
Normal	0.003	0.128	497.2	0.031	0.142	500.4	0.001	0.151	508.6	0.007	0.139	532.3			
Log-normal3	0.021	0.087	495.5	0.123	0.091	505.3	0.012	0.091	506.4	0.212	0.095	513.9			
Log-Pearson 3	0.049	0.097	499.0	0.114	0.101	508.4	0.014	0.092	507.1	0.474	0.078	508.2			
Gumbel	**0.082**	0.083	**491.7**	0.196	0.093	**499.5**	**0.027**	**0.088**	**502.2**	0.485	**0.085**	**507.6**			
Log-Logis3	0.058	**0.074**	494.5	**0.256**	**0.071**	501.5	0.010	0.089	505.2	**0.489**	0.090	508.4			

- CHS = Chi –Square Test (P value)
- KS = Kolmogorov-Smirnov test (Statistic)
- BIC = Bayesian Information Criterion
- Bold number representing the best result

Appendix 2. Calculation of frequency intensity I_T (mm/hr) values for different durations t_d (minutes and hours) and return periods *Tr* (years) using Gumbel method**.**

Computed Intensity (I_T) Gumbel Method								
t_d		5 min	10 min	15 min	30 min	1 hr	2 hr	3 hr
	Tr			K_T			I_T	
2	-0.16	108.07	82.40	70.63	51.62	38.24	26.08	20.68
5	0.72	175.34	122.23	102.07	73.37	53.60	35.02	27.87
10	1.30	219.88	148.61	122.88	87.77	63.77	40.94	32.62
25	2.04	276.16	181.93	149.18	105.96	76.62	48.42	38.63
50	2.59	317.91	206.64	168.69	119.45	86.15	53.97	43.09
100	3.14	359.35	231.18	188.06	132.85	95.61	59.48	47.51
500	4.39	455.11	287.88	232.81	163.81	117.48	72.21	57.74
t_d		4 hr	5 hr	6 hr	8 hr	12 hr	18 hr	24 hr
Tr	K_T				I_T			
2	-0.16	17.30	15.32	13.58	11.03	8.17	6.11	5.09
5	0.72	23.05	20.25	17.80	14.63	10.65	7.91	6.79
10	1.30	26.86	23.51	20.59	17.01	12.29	9.10	7.92
25	2.04	31.68	27.63	24.12	20.03	14.36	10.60	9.35
50	2.59	35.25	30.69	26.74	22.26	15.90	11.71	10.41
100	3.14	38.79	33.73	29.34	24.48	17.43	12.82	11.46
500	4.39	46.98	40.74	35.35	29.61	20.96	15.38	13.89