

# Department of Statistics

## COURSE STATS 330

### Assignment 3, 2005

**Instructions:** Hand in your completed assignment to the Student Resource Centre by 4pm on Thursday 15<sup>th</sup> September. Note that this assignment **requires you to write a report**.

You are employed by a firm of statistical consultants, who have been approached by a local council, who need help with a study into water run-off problems arising from severe storms.

The council wants to predict how much peak runoff (measured in cubic meters per second) will follow storms of different intensities. The council has collected data on six different watersheds within its region, and wants predictions on all six.

There are several variables measured for each watershed.

These are

**area:** Area of catchment (sq km)  
**slope:** Average slope of watershed  
**long:** Longest stream flow ( km)  
**infil:** infiltration rate of water into soil (mm/hour)

The values are given in the following table:

watershed	area	slope	long	infiltration rate
1	0.078	3	0.305	6.35
2	0.337	6.5	0.610	8.89
3	2.590	15	3.048	5.08
4	7.770	7	4.572	12.70
5	12.95	6	4.572	15.24
6	18.13	6.5	5.791	12.70

These data are in the file **watershed.csv**.

In addition, the council has data on 30 previous storms. For each of these, the following variables are recorded:

<b>watershed:</b>	The watershed being recorded, one of 1, 2, 3, 4, 5, 6.
<b>imperv:</b>	The area in the watershed impervious to water (sq km). This changes with time, the figure given is the value at the time of the storm.
<b>absorb:</b>	Surface absorbency index (0=no absorbency, 100=total absorbency) This depends on the amount of water in the ground prior to the storm.
<b>storage:</b>	Estimated soil storage capacity (mm of water) This also depends on the amount of water in the ground prior to the storm.
<b>rainfall:</b>	Rainfall in mm during the storm.
<b>intensity:</b>	Time period (hours) for which the rainfall exceeded 5mm/hour.
<b>peak.flow</b>	Peak flow of water (cubic meters/second).

The values of these variables for each of the 30 storms is given in the file **storms.csv**, and is also reproduced overleaf.

Use these two sets of data to develop a set of predictions for each watershed. You are required to predict peak flows for each watershed, for storms having 3 values of rainfall (50, 75 and 100 mm) and 2 values of intensity (2 and 5 hours). For the other variables, use either the watershed figure, or the average of the watershed figures if these differ from storm to storm.

Write a report to the council which presents these predictions. Your report should not exceed 8 pages, and should have an executive summary, an introduction, a middle section and a conclusion. Technical details of model fitting should be placed in an appendix. A feature of the report will be a table giving the predictions, together with some idea of their accuracy.

I have placed a Powerpoint presentation on report writing on the web page, which you may find helpful.

Storm data

watershed	imperv	absorb	storage	rainfall	intensity	peak.flow
1	0.016	70	38.1	44.45	2	46
1	0.016	70	38.1	57.15	3.7	28
1	0.016	70	38.1	101.6	4.2	54
1	0.054	80	25.4	40.64	1.5	70
1	0.054	80	25.4	78.74	4	47
1	0.054	80	25.4	91.44	2.4	112
2	0.013	65	50.8	31.75	0.7	398
2	0.013	65	50.8	58.42	3.5	98
2	0.013	65	50.8	107.95	4	191
2	0.021	68	12.7	36.83	2	171
2	0.021	68	12.7	66.04	4	150
2	0.021	68	12.7	99.06	3	331
3	0.06	60	25.4	19.05	1	772
3	0.06	60	25.4	44.45	1.5	1268
3	0.06	60	25.4	82.55	4	849
3	0.06	65	50.8	45.72	1	2294
3	0.06	65	50.8	78.74	2	1984
3	0.06	65	50.8	120.65	6	900
4	0.101	67	12.7	44.45	2	2181
4	0.101	67	12.7	82.55	4	2484
4	0.101	67	12.7	127	6.5	2450
5	0.282	62	38.1	38.1	1.5	1794
5	0.282	62	38.1	69.85	3	2067
5	0.282	62	38.1	106.68	5	2586
6	0.142	56	50.8	45.72	2	2410
6	0.142	56	50.8	82.55	4	1808
6	0.142	56	50.8	133.35	6	3024
6	0.163	56	25.4	31.75	2	710
6	0.163	56	25.4	73.66	3.4	3181
6	0.163	56	25.4	120.904	5	4279