Chapter 4

Methodology

4.2 Statistical Methods

From provincial fire management records in the Ontario Ministry of Natural Resources (OMNR) fire database the following annual statistics will be calculated:

- E_{t} the annual number of escapes (i.e., fires $\geq$3 *ha*)
- L_{t} the annual number of large fires (i.e., fires $\geq$200 *ha*)

As Cumming (2005:775) suggests, the observed annual distribution of fires (L_{t} / E_{t}) will be used as an estimate of the annual probability ($\rho \mathrm{t}$) that a randomly chosen escaped fire (E) will fail to be suppressed and become a large fire (L).

Conditional on E_{t} (i.e., the annual number of escapes), L_{t} (i.e., the annual number of large fires) will, therefore, be a *binomial random variable* with an expected value of $L_{\mathrm{t}}\ \rho \mathrm{t}$ and variance $L_{\mathrm{t}}\ \rho \mathrm{t}\ (1 - \rho \mathrm{t})$. As such, the hypotheses H_0 and H_{S} can be operationalised as logistic regression models of the annual variation in $\rho \mathrm{t}$, and tested by regressing the observations against the suppression strategy (S_{i} or S_{m}) being employed.

With the OMNR dataset .csv file downloaded to the working directory, the file was loaded into R.

```
> omnr <- read.csv("OMNR_data.csv")
```

To calculate N_{t} the annual load, the total number of fires per year were aggregated. The OMNR dataset was then split into the two Fire Management Zones, Inensive and Measured, reqired for the analysis.
The OMNR dataset was further seperated into the number of suppressed and escaped fires in each Fire Management Zone.

N_s and N_e could then aggregated.

With the following equation

```
glm.eq <- "cbind(omnr.glm$suppressed, omnr.glm$unsuppressed)
omnr.glm$zone * omnr.glm$year"
```

```
> options(show.signif.stars=F)
> glm.eq <- "cbind(omnr.glm$suppressed, omnr.glm$unsuppressed)~ omnr.glm$zone * omnr.glm$yea
> glm.out <- glm(glm.eq, family=binomial(logit),
+                data=omnr.glm)
> summary(glm.out)

Call:
glm(formula = glm.eq, family = binomial(logit), data = omnr.glm)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-8.3516  -0.3858   0.6594   1.8976   5.0376

Coefficients:
                                 Estimate Std. Error z value Pr(>|z|)
(Intercept)                      35.03499   40.31208   0.869   0.3848
omnr.glm$zonemea                146.85587   70.22388   2.091   0.0365
omnr.glm$year                    -0.01650    0.02019  -0.817   0.4138
omnr.glm$zonemea:omnr.glm$year   -0.07409    0.03519  -2.105   0.0353

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 238.46  on 31  degrees of freedom
Residual deviance: 183.76  on 28  degrees of freedom
AIC: 288.11

Number of Fisher Scoring iterations: 5

>
```

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	35.0350	40.3121	0.87	0.3848
omnr.glm$zonemea	146.8559	70.2239	2.09	0.0365
omnr.glm$year	-0.0165	0.0202	-0.82	0.4138
omnr.glm$zonemea:omnr.glm$year	-0.0741	0.0352	-2.11	0.0353

Table 1: Summary statistics for the regression model