11e colloque du cef
Centre d'étude de la forêt

UNIVERSITÉ LAVAL

Modeling the number and size of forest fires in Canada

Carla Francisco [1], Steven G. Cumming[2]

[1,2] Faculté de Foresterie, de Géographie et de Géomatique, Québec, Canada

Introduction

This paper is devoted to a problem of developing statistical models for forecasting the number and size of forest fires in Canada. Fire management and size-biased sampling are likely to have affected the historical record. This is because many small fires were likely undetected, while many detected fires were actioned. This would bias naïve models of the relations between fire weather and the fire regime parameters of interest. It is becoming important to develop more sophisticated models that account for this effect in order to make reliable forecasts under climate change.

Statistical analysis

Below we have a photo of the big fire in Fort McMurray:

Fig.1 - Fort McMurray fire, Canada. Foto: @jeromegarot/Twitter.

In Fig.2 on the left side , we have the representation of the number of fires by year, between 1950 and 2015. On the right side in Fig.3 we have a representation of the number of fires by year:

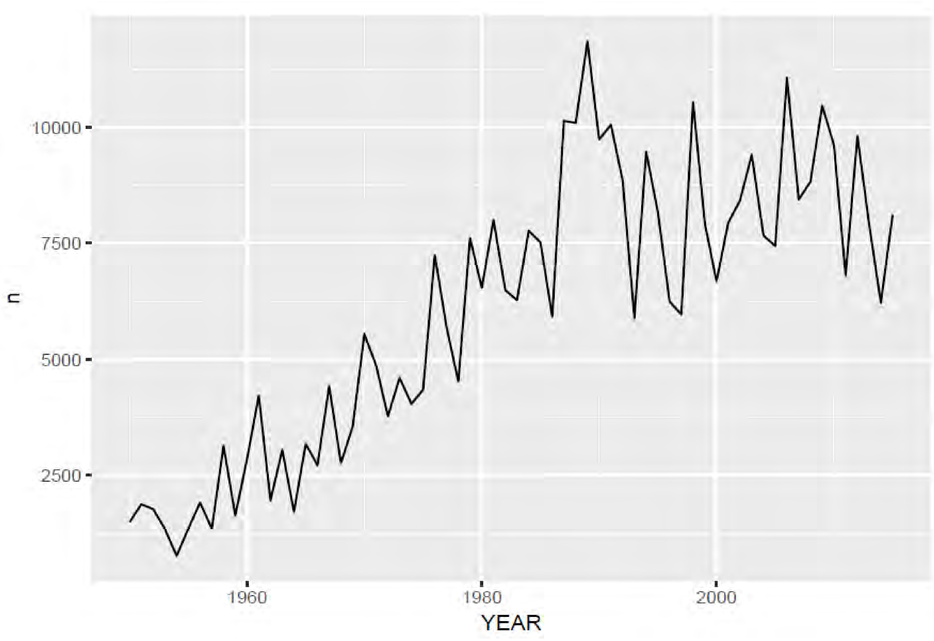


Fig.2 - Evolution fires vs year.

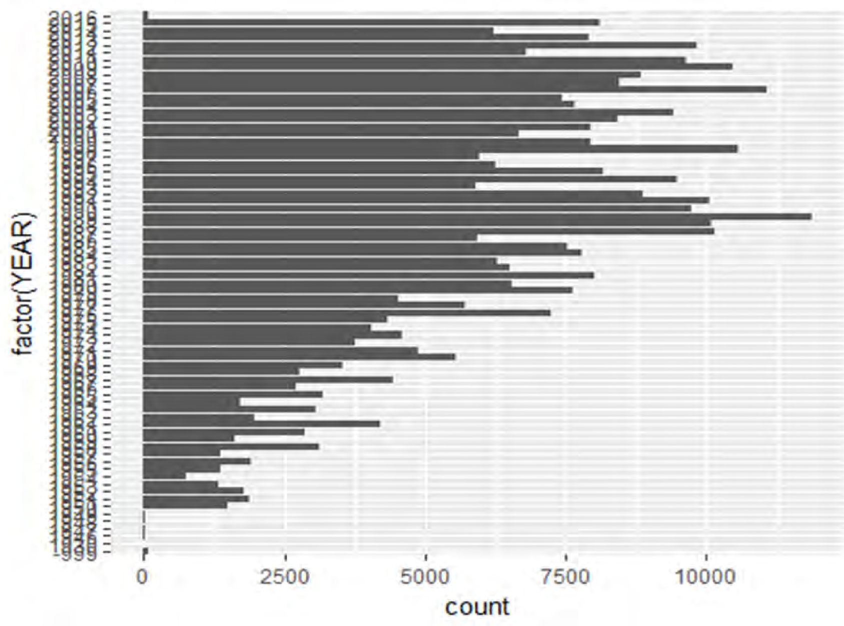


Fig.3 - Graph of the number of fires by year.

Below on Fig.4, we have the representation of the evolution of the mean size of fire vs year and on Fig.5 we have the relationship between number of fires(n) and mean size of fire (mean size).

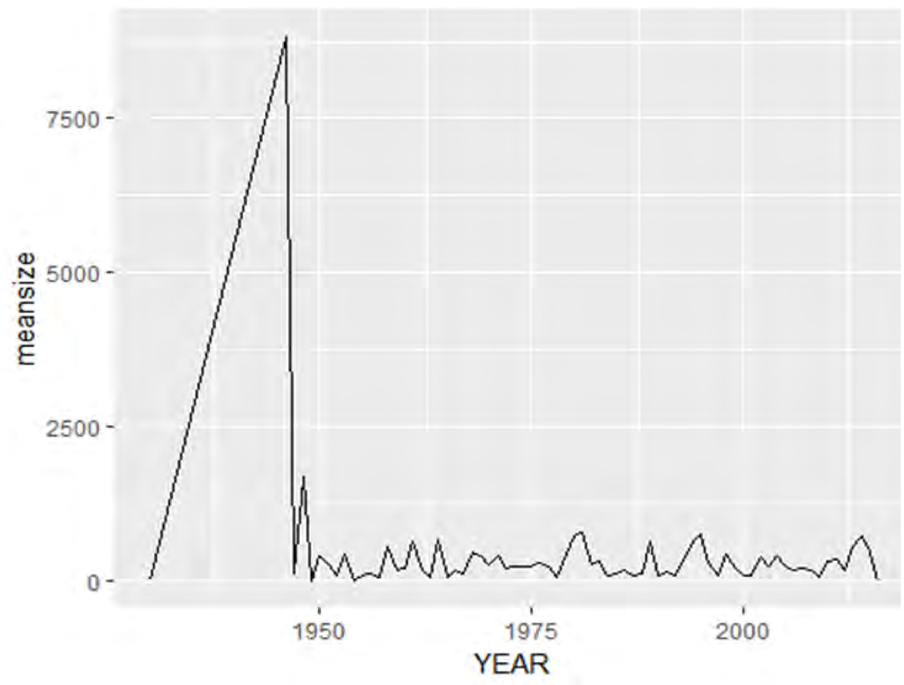


Fig.4 - Evolution of the mean size of fire vs year.

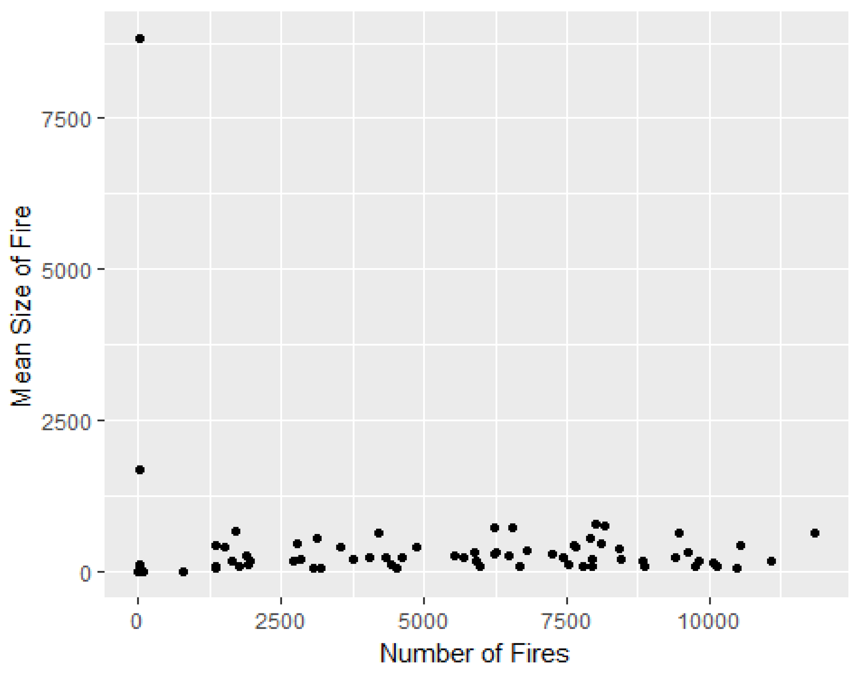
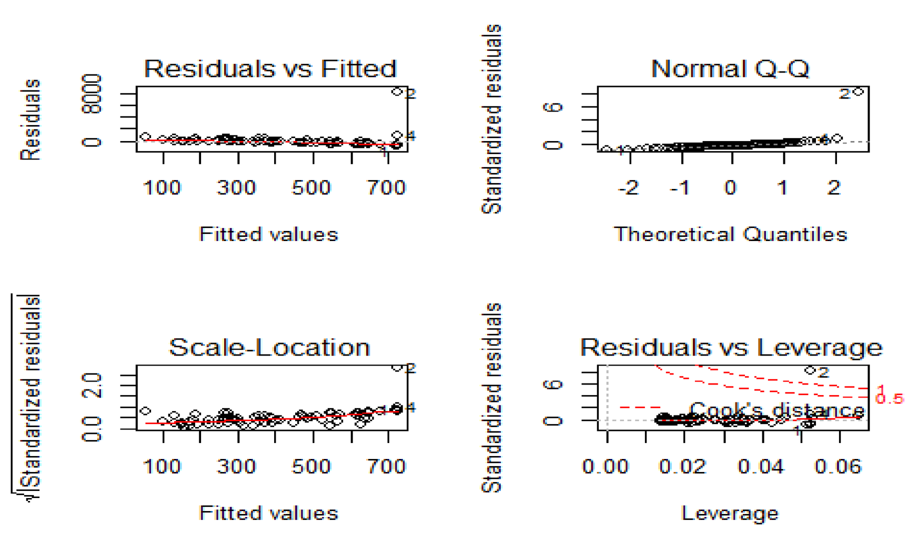


Fig.5–Evolution of number of fires vs mean size of fire.

The model is little significant for variable n. We can conclude that for the whole data there is no relation between the number of fires and the mean size of the fire. The intercept has a high significance (0.0001), but the coefficient of the independent variable(n) is no significant. Analysis of regression of the total data:



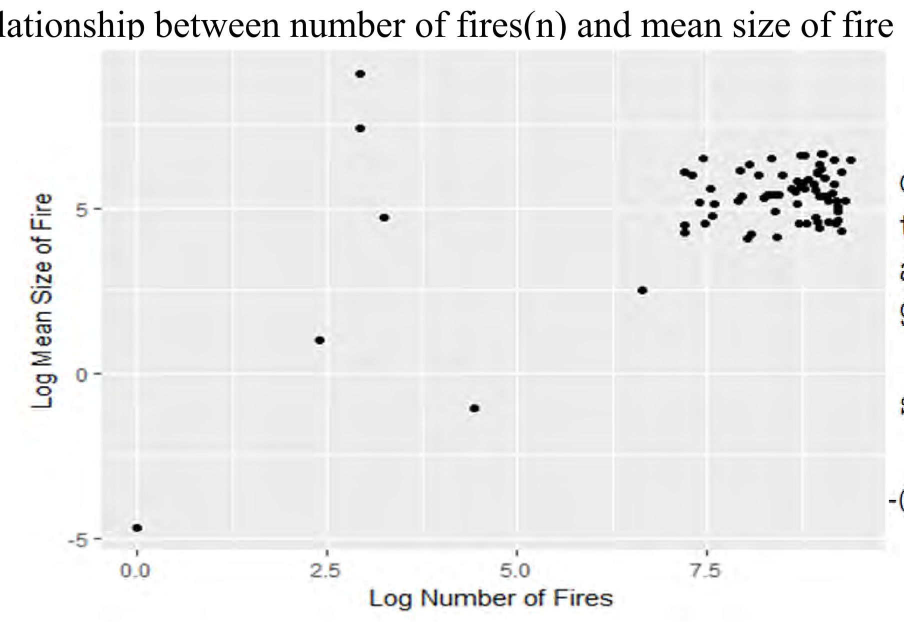

```
Residuals:
   Min     1Q Median     3Q    Max
-722.2 -319.5  -86.6   37.6 8103.8

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 722.28887  236.41241   3.055  0.00318 **
n            -0.05611    0.03673  -1.528  0.13108
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1030 on 70 degrees of freedom
Multiple R-squared:  0.03227,    Adjusted R-squared:  0.01844
F-statistic: 2.334 on 1 and 70 DF,  p-value: 0.1311
```

Fig.6 - Representation of the Analysis of regression of the total data.

Relationship between number of fires(n) and mean size of fire (mean size), using logarithmic scales:

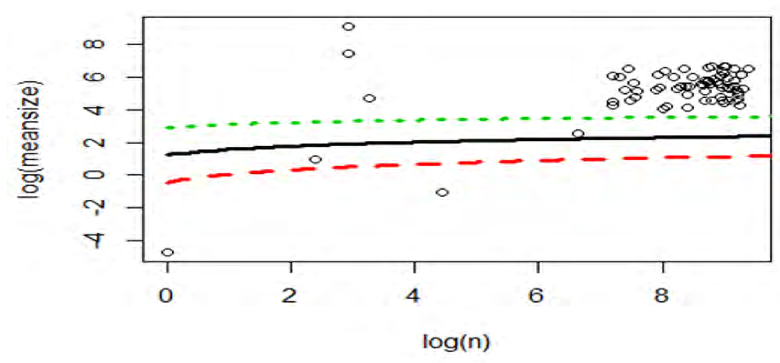



```
Pearson's product-moment correlation

data:  data5$n and data5$meansize
t = -1.5278, df = 70, p-value = 0.1311
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 -0.39486794  0.05429627
sample estimates:
       cor
-0.1796318
```

Fig.7- Number of fires(n) vs mean size of fire (mean size), using logarithmic scales.

Representation of the predict model:




```
Residuals:
    Min      1Q  Median      3Q     Max
-5.9283 -0.5944  0.0711  0.6055  6.4194

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   1.2178     0.8361   1.457     0.15
log(n)        0.4918     0.1015   4.844 7.37e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 1.535 on 70 degrees of freedom
Multiple R-squared:  0.251,  Adjusted R-squared:  0.2403
F-statistic: 23.46 on 1 and 70 DF,  p-value: 7.371e-06
```

Fig.8 – Graph of the of the predict model.

In Fig.9 on the left side , we have the representation with Outliers. On the right side in Fig.10 we have a representation without Outliers:

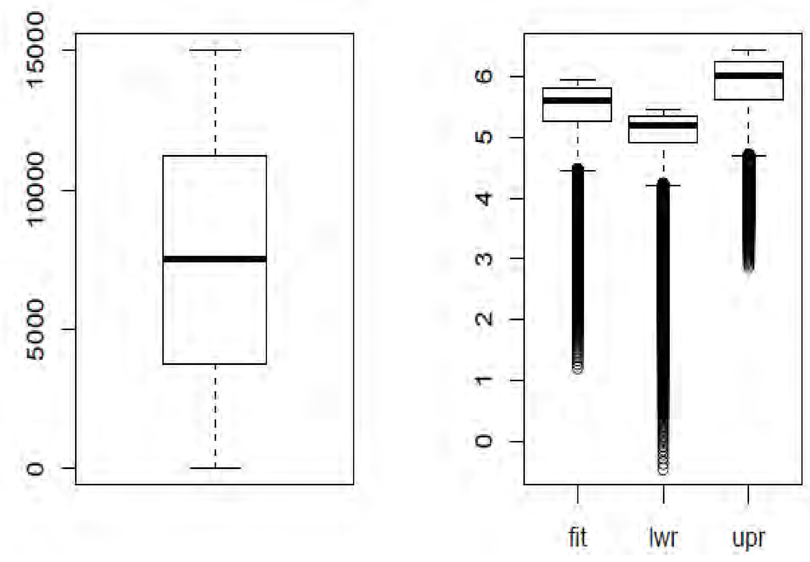


Fig.9 - Representation with Outliers.

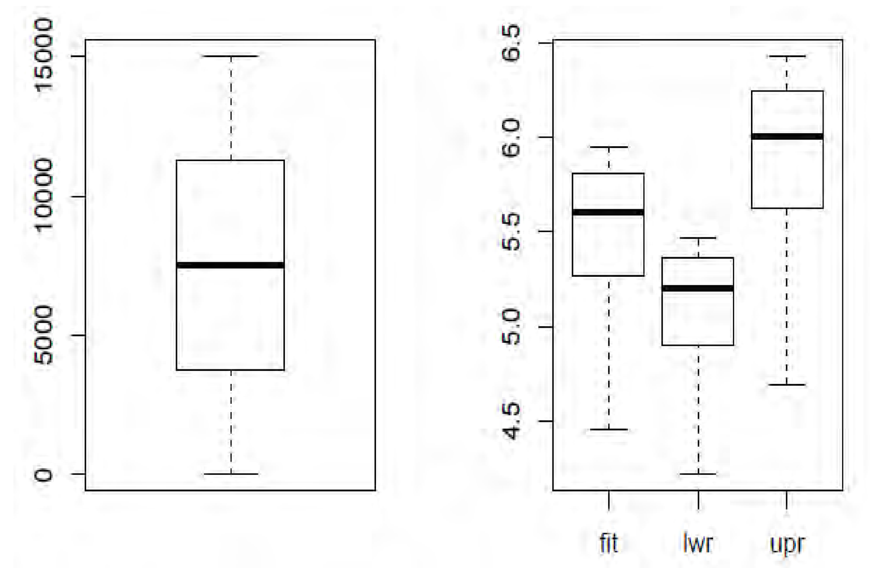


Fig.10 - Representation without Outliers.

In Fig.11 on the left side we have the representation of the regression model with Outliers. For the threshold we pick one that produces 24 Outliers, if we choose a lower number, the Outliers increase, for example, for threshold = 0.5 there are 39 Outliers. On the right side in Fig.12 we have the regression representation without Outliers:

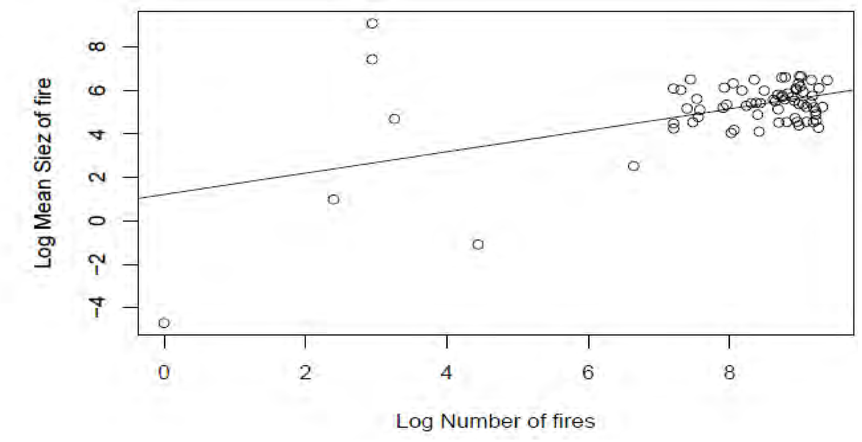


Fig.11 - Regression representation with Outliers.

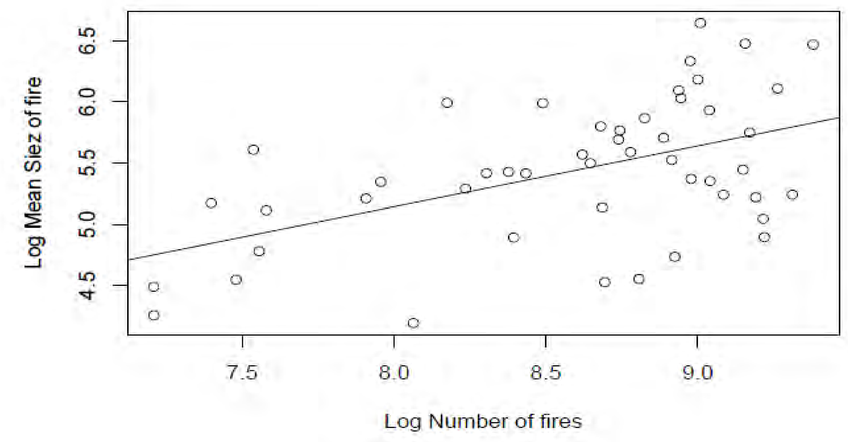
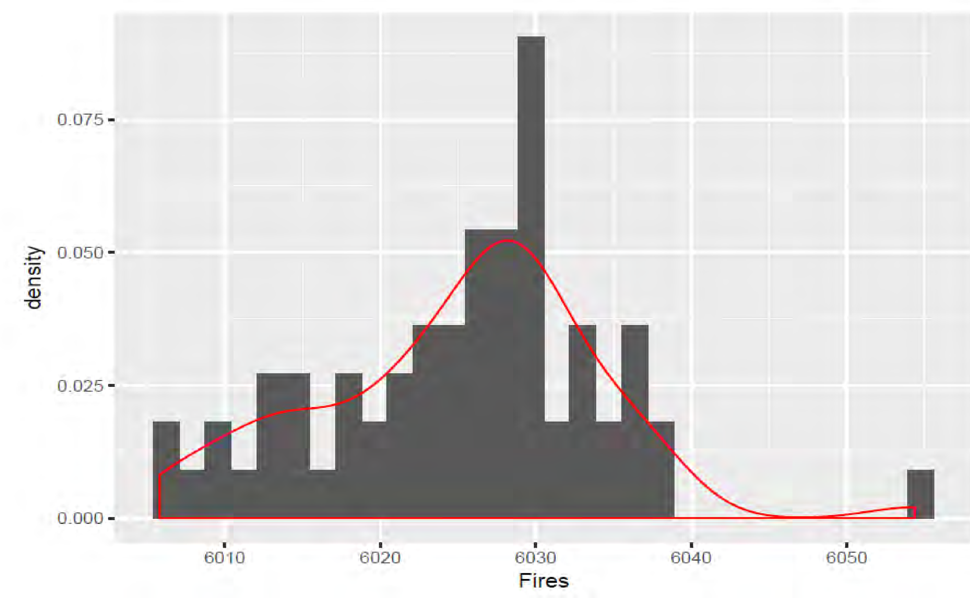


Fig.12 - Regression representation without Outliers.

We created a matrix with 100000 rows , one for each Bootstrap sample and 72 columns, one for each sampled values, to match the original sample size, and we obtained the mean value of the number of fires for each year, the graph for density distribution of number of fires, and the endpoints for 90%, 95% and 99% Bootstrap confidence intervals using percentiles.




```
quantile(boot.statistics,c(0.05,0.95))

##       5%      95%
## 6009.150 6036.938

quantile(boot.statistics,c(0.025,0.975))

##     2.5%    97.5%
## 6007.565 6038.593

quantile(boot.statistics,c(0.005,0.995))

##     0.5%    99.5%
## 6006.184 6049.296
```

Fig.13 – Representation of the Bootstrap of the total data.

Conclusions

With these temporal dataset of the number of fires we calculated a distribution which approaches the fire density within the mean, the variance and skewness of the proposed distribution, in order to model the number and size of forest fires. We can conclude that for the whole data there is no relation between the number of fires and the mean size of the fire.

References

1. A leisurely look at the bootstrap, the jackknife, and cross-validation. B Efron, G Gong - The American Statistician, 1983.
2. An introduction to the bootstrap. B Efron, RJ Tibshirani, 1994.
3. Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. B Efron, R Tibshirani - Statistical science, 1986.
4. Bootstrap confidence intervals for a class of parametric problems. B Efron - Biometrika, 1985.
5. Computer intensive methods in statistics. P Diaconis, B Efron, 1983.
6. The Latex companion. Michel Goossens, Frank Mittelbach, and Alexander Samarin. AddisonWesley, Reading, Massachusetts,1993.
7. More efficient bootstrap computations. Efron, Bradley. Journal of the American Statistical Association 85.409, 1990: 79-89.
8. Natural Resources Canada - Northern Forestry Centre and Canadian National FireDatabase, Accessed in March, 2017. Available in: http://www.nrcan.gc.ca/forests/research-centres/nofc/13485
9. R Core Team. R: A Language and Environment for Statistical Computing. Vienna: R Foundation for Statistical Computing, 2013. Available: http://www.R-project.org.
10. The jackknife estimate of variance. B Efron, C Stein - The Annals of Statistics, 1981.
11. The jackknife, the bootstrap and other resampling plans. B Efron, 1982.