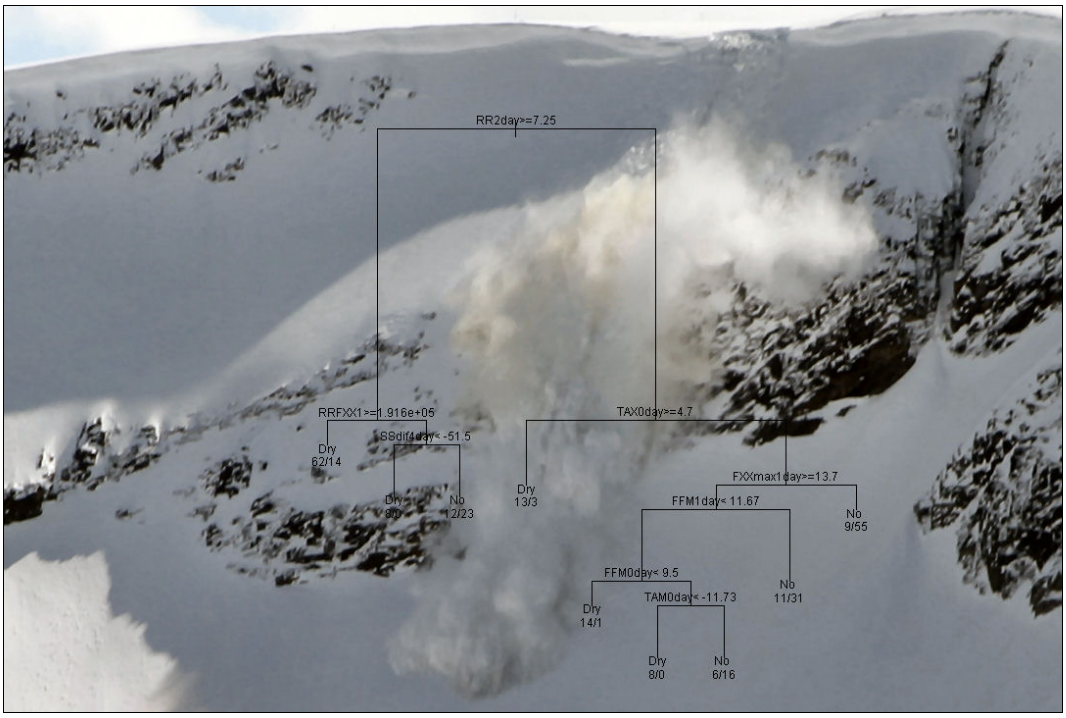
Master Thesis in Geosciences

Snow Avalanche Prediction in Grasdalen, Norway

Application of wind drift factors and classification trees

Emma Kate Howley



UNIVERSITY OF OSLO

FACULTY OF MATHEMATICS AND NATURAL SCIENCES

Snow Avalanche Prediction in Grasdalen, Norway

Application of wind drift factors and classification trees

Emma Kate Howley

Master Thesis in Geosciences
Discipline: Environmental Geology and Geohazards
Department of Geosciences
Faculty of Mathematics and Natural Sciences

UNIVERSITY OF OSLO

June 2007

Tutors: Professor Kaare Høeg (UiO) and Dr. Kalle Kronholm (NGI)

Cover photo: Norwegian Geotechnical Institute (NGI)

Acknowledgements

Firstly I would like to take this opportunity to thank Dr. Kalle Kronholm at NGI who has done a fantastic job as supervisor in advising, and encouraging me throughout the period of my thesis study. More specifically, thank you for all your help and guidance with the probability plots, and also the R script, without which this thesis would not have got off the ground! I would also like to say a big thank you to Professor Kaare Høeg for both helping with the early stages of my topic choice and also for all his support and words of wisdom during the last 2 years of my studying at UiO, I really enjoyed all our modules together.

I would finally like to acknowledge and thank my family and friends for their support and encouragement, in particular my Mum and Dad, thank you for all your advice and support over the years. And to Mum and Ruth - thank you for proof reading and spell checking where the computer failed!

Emma Kate Howley
Oslo, June 2007

Abstract

Snow avalanches are a common hazard experienced in many mountainous environments around the world during the winter months. It is therefore often of great importance to be able to predict these events to reduce the risk they pose to the population and infrastructure. Work done in this thesis has involved detailed analysis of various weather parameters within Stryn in western Norway in order to establish common triggering factors for avalanches in Grasdalen. Two large sets of data; a gridded extrapolated data set and an observed set of data from the study area have been compared. From these data sets, several wind drift factors have been derived as these parameters are considered important avalanche triggering elements by several authors. These combined factors have not previously been analysed for the observed Fonnbu data set. Statistical procedures include cumulative probability plots which have provided threshold values, a Kruskal-Wallis test, and additionally, a number of classification trees. The latter were used to highlight the most important weather parameters used to classify data in terms of dry avalanche days or non avalanche days which has not been undertaken with the Fonnbu data set previously. Results of these indicate the primary splitting factor to be various sums of precipitation over the preceding days, particularly the two day sum and four day sum. Following this, the maximum temperature measured on the preceding day is considered important for classification tree splits between dry avalanche and non avalanche days. In terms of the combined wind drift factors, these appear in 32 % of the classification trees within the top three splits, the most predominant being calculated using maximum wind speeds on the day of the avalanche. Although this thesis is not in total agreement with the results of previous work which outlines the importance of the wind drift factor, rather, it highlights the complicated relationship between preceding weather conditions and avalanche occurrence indicating the vast array of factors to be considered for avalanche prediction.

Keywords: snow avalanche, forecasting, prediction, meteorology, wind drift, classification tree, Norway.

Table of Contents

Glossary of Terms

GLOSSARY OF TERMS FOR GRIDDED DATA		
No	avalanche event 1-886	
slide_ID	avalanche name 1-1018	
Codename	rank 1-51	
Slide_Date Accuracy	+/- time error	hours/mins
Global_Exposition Code	aspect of avalanche path	direction
Avalanche_Snow WetnessCode	wet/dry/unknown	
rr1day	precipitation on day of avalanche	mm
rr3day	three day sum of precipitation ending at the end of the avalanche day	mm
rr5day	five day sum of precipitation ending at the end of the avalanche day	mm
Tam	daily mean air temperature	˚C
wndspd1day	average wind speed on the day of the avalanche	m/s
wnddir1day	direction of average wind speed on the day of the avalanche	degrees
sector1day	wnddir1day given 1 of 8 catergories N,NE,E,SE,S,SW,W,NW	direction
Wndspdmax1day	maximum wind speed on the day of the avalanche	m/s
wnddirmax1day	direction of maximum wind speed on the day of the avalanche	degrees
sector1dmax	winddirmax1day given 1 of 8 categories N,NE,E,SE,S,SW,W,NW	direction
wndspd3day	three day average wind speed ending at the end of the avalanche day	m/s
wnddir3day	direction of wndspd3day	degrees
sector3day	winddir3day given 1 of 8 categories N,NE,E,SE,S,SW,W,NW	direction
Wndspdmax3day	three day maximum wind speed ending at the end of the avalanche day	m/s
wnddirmax3day	direction of wndspdmax3day	degrees
sector3dmax	winddirmax3day given 1 of 8 categories N,NE,E,SE,S,SW,W,NW	direction
wndspd5day	five day average wind speed ending at the end of the avalanche day	m/s
wnddir5day	direction of wndspd5day	degrees
sector5day	winddir5day given 1 of 8 categories N,NE,E,SE,S,SW,W,NW	direction
Wndspdmax5day	five day maximum wind speed ending at the end of the avalanche day	m/s
wnddirmax5day	direction of wndspdmax5day	degrees
sector5dmax	winddirmax5day given 1 of 8 categories N,NE,E,SE,S,SW,W,NW	direction
COMBINED PARAMETERS:		
rrwndspd1	rr1day x (wndspd1day)^4	mm(m/s)^4
rrwndspd3	rr3dayx (wndspd3day)^4	mm(m/s)^4
rrwndspd5	rr5day x (wndspd5day)^4	mm(m/s)^4
rrwndspdmax1	rr1day x (wndspdmax1day)^4	mm(m/s)^4
rrwndspdmax3	rr3day x (wndspdmax3day)^4	mm(m/s)^4
rrwndspdmax5	rr5day x (wndspdmax5day)^4	mm(m/s)^4

GLOSSARY OF TERMS FOR FONNBU DATA		
n.metNAs	number of unknown values on each day	
Allmet	false if unknowns, true if all values known	
nAval	total number of avalanches each day	
nAval.dry	number of dry avalanches each day	
nAval.wet	number of wet avalanches each day	
nAval.mixed	number of mixed avalanches each day	
nAval.unknown	number of avalanches of unknown type each day	
RR0day	precipitation on the preceding day	mm
RR1day	precipitation on the day	mm
RR2day	two day sum of precipitation	mm
RR3day	three day sum of precipitation	mm
RR4day	four day sum of precipitation	mm
RR5day	five day sum of precipitation	mm
TAM0day	mean temperature on the preceding day	˚C
TAM1day	mean temperature on the day	˚C
TAN0day	minimum temperature on the preceding day	˚C
TAN1day	minimum temperature on the day	˚C
TAX0day	maximum temperature on the preceding day	˚C
TAX1day	maximum temperature on the day	˚C
SS0day	snowdepth on the preceding day	cm
SS1day	snowdepth on the day	cm
SSdif1day	depth of new snow on the day	cm
SSdif2day	two day sum of depth of new snow	cm
SSdif3day	three day sum of depth of new snow	cm
SSdif4day	four day sum of depth of new snow	cm
SSdif5day	five day sum of depth of new snow	cm
FFM0day	daily mean wind speed at 10m above the ground (averaged over 1-hourly measurements) on the preceding day	m/s
FFM1day	daily mean wind speed at 10m above the ground (averaged over 1-hourly measurements) on the day	m/s
FFX0day	maximum mean wind speed on preceding day	m/s
FFX1day	maximum mean wind speed on day	m/s
FXM0day	mean maximum wind speed on preceding day	m/s
FXM1day	mean maximum wind speed on day	m/s
FXX0day	highest maximum wind speed on preceding day	m/s
FXX1day	highest maximum wind speed on day	m/s
FXXmax1day	maximum highest maximum wind speed recorded over the one day period	m/s
FXXmax2day	maximum highest maximum wind speed recorded over the two day period ending on the day	m/s
FXXmax3day	maximum highest maximum wind speed recorded over the three day period ending on the day	m/s
FXXmax4day	maximum highest maximum wind speed recorded over the four day period ending on the day	m/s
FXXmax5day	maximum highest maximum wind speed recorded over the five day period ending on the day	m/s

COMBINED PARAMETERS:		
RRFFM0	RR0day x $(FFM0day)^4$	$mm(m/s)^4$
RRFFM1	RR1day x $(FFM1day)^4$	$mm(m/s)^4$
RRFFX0	RR0day x $(FFX0day)^4$	$mm(m/s)^4$
RRFFX1	RR1day x $(FFX1day)^4$	$mm(m/s)^4$
RRFXM0	RR0day x $(FXM0day)^4$	$mm(m/s)^4$
RRFXM1	RR1day x $(FXM1day)^4$	$mm(m/s)^4$
RRFXX0	RR0day x $(FXX0day)^4$	$mm(m/s)^4$
RRFXX1	RR1day x $(FXX1day)^4$	$mm(m/s)^4$
RRFXXmax1	RR1day x $(FXXmax1day)^4$	$mm(m/s)^4$
RRFXXmax2	RR2day x $(FXXmax2day)^4$	$mm(m/s)^4$
RRFXXmax3	RR3day x $(FXXmax3day)^4$	$mm(m/s)^4$
RRFXXmax4	RR4day x $(FXXmax4day)^4$	$mm(m/s)^4$
RRFXXmax5	RR5day x $(FXXmax5day)^4$	$mm(m/s)^4$

Chapter 1: Introduction

1.1 Aim

Meteorological data with particular reference to precipitation, wind speed and combined wind drift parameters from the Stryn district in western Norway will be analysed. Statistical methods will be used to distinguish between days with and days without avalanches with a view to aid avalanche prediction.

1.2 Background

Around the world, as the population increases and people become increasingly affluent, enjoying more leisure time, there is a growing trend to encroach on more remote and fragile environments. Infrastructure has expanded to support this, with road and rail networks traversing hazardous terrain to link communities. In Canada and the U.S. numerous transport corridors pass through renowned avalanche paths. In New Zealand the Milford Road is frequently affected by avalanche hazards (Fitzharris *et al*., 1999; Hendrikx *et al*., 2005), and in Norway Kristensen *et al*. (2003) report that avalanches cause 70 to 80 % of all road blockages in the country. In addition, the recent boom of the ski and outdoor activity industry seen within the European Alps has led to more frequent reporting and monitoring of the avalanche hazard over the last 50 years. The winter season of 1999 saw the devastating impact in ski resorts across the Austrian, Swiss and French Alps as numerous avalanches killed inhabitants and tourists, caused untold damage and left many snowbound without necessary supplies. The cause of these 1999 avalanches was frequent and heavy snowfall accompanied by high winds (RTD, 2006).

The above are just a few examples to outline the growing necessity to be able to predict avalanche occurrence. In many avalanche prone areas implementation of hazard or risk maps with the use of GIS techniques aids in outlining known avalanche paths and appropriate land use zoning (Gruber and Margreth, 2001; Walsh *et al*., 1990; Furdada *et al*., 1995). In addition, several methods in snow stability testing have been applied in relation to avalanche prediction and formation (Birkeland *et al*., 1996; McElwaine *et al*., 2000; Birkeland, 2001; Landry *et al*., 2002). However, one of the most useful methods of avalanche prediction is that of monitoring and analysing preceding weather conditions.

1.3 Purpose of study

Several investigations have been carried out in order to quantify links between certain weather parameters and avalanche occurrence. Bakkehøi (1987) used a probability distribution method, this established a good correlation using the three day sum of precipitation, to predict avalanche occurrence on several release paths in Stryn, western Norway. In the same area, Kronholm *et al*. (2006a; 2006b) have recently looked at the role of classification trees using extrapolated gridded data sets in which the one and five day sums of precipitation were of significance for avalanche day prediction. Further work has additionally been carried out using classification trees created by Davis *et al*. (1999) for areas of Utah and California, and by Hendrikx *et al*. (2005) for the Milford Road, New Zealand. These detail the high rate of correct prediction by distinguishing between avalanche and non avalanche days when combining precipitation and wind speed into a wind drift parameter.

In this investigation, similar concepts of Hendrikx *et al*. (2005) and Davis *et al*. (1999) will be implemented by combining data of different weather parameters, more specifically, precipitation and wind speed into a wind drift parameter. This is with particular reference to data from the Fonnbu weather station near Stryn, western Norway. This data set has not previously been used in any avalanche classification procedures to distinguish between avalanche days and non avalanche days. Classification trees implementing the combined wind drift data and other weather parameters will be created, and certain threshold parameter values will be summarised. It is therefore hoped that similar results to Hendikx *et al*. (2005) and Davis *et al*. (1999) regarding avalanche release probability can be established to aid with prediction and decision making within the area surrounding Grasdalen in Stryn.

Chapter 2: Literature Survey

An avalanche is a generic term used to describe a falling mass of either one or more of the following; snow, ice, rock and/or debris under the influence of gravity (McClung and Schaerer, 1993). According to Perla (1980) avalanches may vary from a *"harmless trickle of loose snow"* sliding to a new angle of repose, to a *"large devastating mass of snow, ice and earth"* which can travel down extensive slopes with great speed and energy. In addition, snow avalanches can be categorised as either loose snow or slab avalanches, which are moreover subdivided into wet or dry avalanches (McClung and Schaerer, 1993).

For clarification, the term avalanche can refer to the fast flowing movement of debris, rock or snow, however this thesis considers snow avalanches only, therefore for convenience snow avalanches shall be read purely as avalanches. These occur where weather conditions permit snow to accumulate on steep enough slopes. Hence, important factors to consider include current and preceding weather conditions and their impact and interaction with the snowpack, and additionally the underlying terrain and topographical attributes at these locations. The above circumstances aid with the process of avalanche prediction in numerous mountainous locations around the world.

2.1 Avalanche path characteristics

Avalanches consist of several sections, these include the starting zone where the unstable snow fails, in the case of loose snow avalanches this is a small point usually within the surface layers of the snow. A slab avalanche on the other hand often begins with a large fracture zone extending to some depth within the snowpack. The track is the slope over which the avalanche moves, this can be an open slope over which the avalanche spreads usually following the fall-line, or a gully in which the avalanche is channelled. The runout zone is defined where the snow decelerates and collects when movement ceases. All these combine as the path which is defined as the fixed locality within which known avalanches move (McClung and Schaerer, 1993).

2.1.1 Topography

Slope angle and curvature are considered highly significant for avalanche formation and are the only factors constant over time. The obvious requirement for initiation is a slope on which snow can accumulate (Schweizer *et al*. 2003). According to McClung and Schaerer (1993) there are no ideal upper and lower bounds for slope angle as these often vary greatly depending on precise location and conditions, however guidelines exist and are presented in table 2.1. It is evident that the preferred inclination for avalanche release is between about 30˚ and 60˚ in the starting zone. Once the avalanche is initiated track inclination can reduce to 15˚, but on reaching a slope less than this, avalanches decelerate quickly coming to rest in the runout zone.

Angle	Description
10˚-25˚	infrequent wet snow avalanches and slush flows
25˚-35˚	infrequent (but large) slab avalanches, wet loose snow
35˚-45˚	slab avalanches of all sizes
45˚-55˚	frequent small slab avalanches
30˚-60˚	dry loose snow avalanches
60˚-90˚	avalanches are rare, small snow sluffs possible

Table 2.1: Starting zone slope guidelines, adapted from McClung and Schaerer (1993).

Regarding topographic characteristics of avalanche starting zones, McClung and Schaerer (1993) state that increased instability is apparent on convex slopes, and a more detailed study by Maggioni and Gruber (2002) established that a concave cross-slope curvature increases avalanche frequency although there can be great variation.

2.1.2 Vegetation

Avalanche paths are often characterised by a lack of well developed vegetation or scar in the vegetation cover of the area. In terms of the starting zone, forests and large stances of vegetation inhibit avalanche formation by intercepting snowfall, and reducing the rate of wind transported snow, in addition the canopy helps regulate the amount of incoming and outgoing radiation limiting the formation of weak snow layers. Finally, trees may act as anchors to help stabilise the surrounding snowpack; McClung and Schaerer (1993) suggest, however, that some smaller shrubs and bushes of willow and alder may exacerbate avalanche conditions as they "*inhibit snow settlement, creating a loose weak* *base for future snowfalls*". Continuing along this vein of thought it is also suggested by McClung and Schaerer (1993) that some avalanches can initiate where long un-cut meadow grass exists compared to short brush-like grass as the longer grass bends over with the weight of the over lying snow and creates a suitable sliding surface.

Unfortunately, once an avalanche has started vegetation cover does little to protect the down-slope area (Bartelt and Stöckli, 2001; McClung and Schaerer, 1993). Large avalanches can break trees and these can then become entrained within the moving debris increasing momentum and causing further damage.

2.2 Meteorological parameters for avalanche triggering

Meteorological circumstances may be considered the most important and also the most variable aspect in the formation of avalanches. Obviously, without low enough temperatures and precipitation there would be no snowfall, this direct snow-loading is considered one of the most likely triggers of avalanches occurring frequently in storm conditions (Perla, 1978). In addition to this direct cause of snow-loading, the weather greatly influences the snowpack stratigraphy creating both strong and weak layers (Logan, 1998). This is highlighted by Butler (1986) who states that strong correlations have been established between meteorological data, snowpack stratigraphy and the occurrence of avalanches, where studies have been undertaken in a variety of locations. In Glacier National Park, Montana, avalanches are associated with several meteorological conditions including "*heavy snow; heavy snows followed by a rise in air temperature to above freezing; a rise in air temperature to above freezing, without precipitation; and rain in association with above-freezing air temperatures*" (Butler, 1986).

It has been shown that a variety of meteorological factors affect avalanche formation in New Zealand. These include not only the general climatic conditions within its mountains of wet, warm and windy weather (Fitzharris *et al.*, 1999), but also more specific causes including heavy snowfall, fluctuating temperatures and frequent periods of rainfall at high elevations (Owens and Weir, 1992 in Fitzharris *et al.*, 1999). The Cairngorms in

Scotland also experience this close relationship between avalanche formation due to heavy snowfall and fluctuating freezing levels. Additionally strong winds associated with storms have great impact in the majority of cases in this mountain range (SAIS, 1999).

In Norway, Bakkehøi (1987) has been able to use the three day sum of precipitation near Stryn, to aid prediction of certain avalanche paths. However, due to the often large variability in threshold limits it is suggested that other important factors to include are wind and temperature along with current knowledge of the local snow stratigraphy.

The above are just a few examples which outline the main preceding weather conditions expected to contribute to the triggering of avalanches at various locations. These will be expanded on along with other weather parameters yet to be mentioned including precipitation intensity and duration, wind direction and speed, sensible heat, and radiation heating or cooling on the snow (McClung and Schaerer, 1993). Although it is difficult to separate the different weather parameters in a review of avalanche forming phenomena, this will be attempted below so each parameter can be discussed in greater detail. Some overlap may occur as; for example snow loading can refer to direct snowfall but also redistribution by wind.

2.2.1 Direct snow-loading

Firstly it is important to note that different authors cite snow depth using varying measurements. In order for any comparisons to be made it is important to differentiate between these measurements in the literature which refer to either; the sum of daily snowfall (daily new-snow increments); the settled depth of a new snow layer accumulated over several days; or the increase of total snow depth (UNESCO, 1981). Additionally, falling snow is described as precipitation in mm water equivalent and the relationship between this and snow depth is outlined in the following equation [Eq. 2.1] which uses measures of density for calculation.

$$\text{Water equivalent (mm)} = \frac{\text{snow depth (mm)} \times \text{snow density (kg/m}^3\text{)}}{\text{Density of water (1000 kg/m}^3\text{)}} \qquad \text{[Eq. 2.1]}$$

It is often the case that as snow thickness increases so does the probability of avalanching (Akkouratov, 1965; de Quervain, 1965; Bakkehøi, 1987). An example from Stryn is presented in figure 2.1. Here Bakkehøi (1987) uses this graphic to predict that precipitation of 45 to 60 mm over a three day period will result in a 50 % probability of avalanche occurrence on the specified paths.

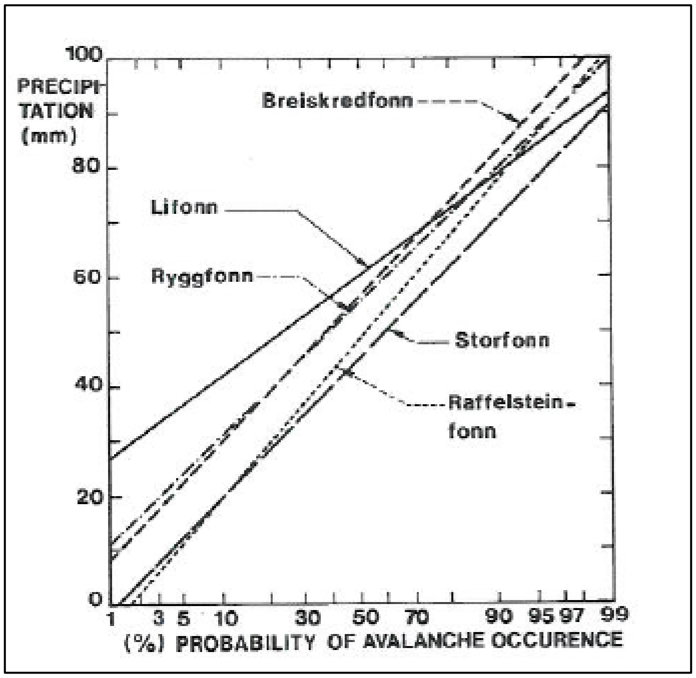


Figure 2.1: Curves represented on normal distribution paper to indicate the probability of avalanche occurrence on five frequently occurring paths with relation to the three day sum of precipitation. (For note Lifonn is also known as Sætreskarsfjellet). Source: Bakkehøi (1987).

Within the same study area Kronholm *et al*. (2006a; 2006b) have applied a classification tree method for avalanche prediction. This used an extrapolated gridded data set as opposed to an observed data set used by Bakkehøi (1987). Although the analysis by Kronholm *et al*. (2006a; 2006b) used a combination of different weather parameters available, it showed one of the most important factors for avalanche prediction to be the five day sum of precipitation. This again highlights the strong influence of preceding snowfall events on avalanche occurrence.

In addition, Zingg (1965) studied the effect of snowfall on avalanche occurrence near Davos in Switzerland, and states that from research during an eight year period, increased snow-load is responsible for most avalanche events. The report goes as far as to say that 69 % of avalanches are highly correlated with either new snow and/or wind transported snow. More recently, from personal qualitative observations in the Haute Savoie department in France during February 1999, an unusually high amount of snowfall over a short period was the major instigator for the large number of avalanches seen in the area. Average snow depths at this time of year of 2 to 3 m of settled snow were exceeded by up to 2 to 3 m of new snow in just a few days. New snow depth is therefore stated to be the most important factor in avalanche warning by UNESCO (1981).

Snowfall / precipitation intensity plays a significant role with regards to snow stability. It is a measure of the rate of increase in depth of snow or mm of water equivalent delivered to the snow per unit time. According to McClung and Schaerer (1993) the measured intensity "*governs the outcome of the race between the shear stress and the increase of strength*" and hence the changes in stability of the snow cover. Several authors have attempted to implement a rule of thumb to relate snowfall / precipitation intensity to the manifestation of instabilities, a conglomeration of these are provided below in table 2.2.

Snowfall intensity (depth per hour cm/hr)	**Precipitation intensity (depth per hour mm/hr)**	**Reference**
> 2.5 cm/hr	0.5 - 2.5 mm/hr	McClung and Schaerer (1993)
1.2 cm/hr - 2 cm/hr	> 2.5 mm/hr	Custer (2005)
~ 1.5 cm/hr	Unknown	La Chapelle (1961) in de Quervain (1965)

Table 2.2: Snowfall and precipitation intensity as a threshold for snow instability.

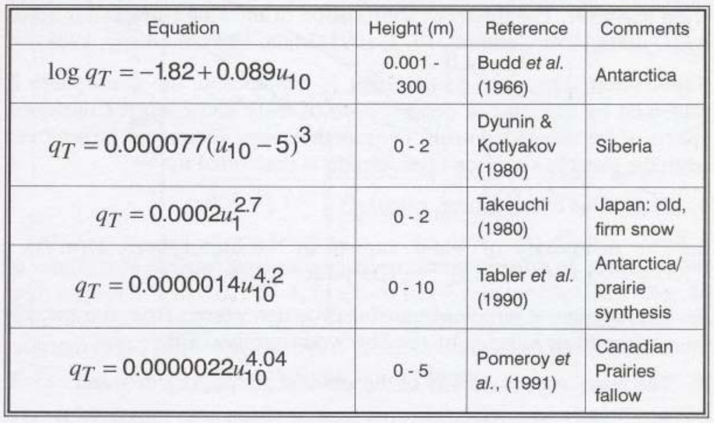
Great variability can be seen between the intensity rates favourable for avalanche formation suggested by the different authors. These differences may be accounted for due to their dependency upon such factors as temperature, wind loading, sluff loading and snow stratigraphy (Custer, 2005; McClung and Schaerer, 1993).

2.2.2 Wind speed

According to de Quervain (1965) after fresh snowfall, wind effects have the next highest impact for the creation of avalanches. During or following snowfall the wind influences the deposited snow creating an irregular and brittle structure (de Quervain, 1965; UNESCO, 1981; Pomeroy and Gray, 1995). Snow drifting occurs causing the snow to be re-distributed, and accumulation is concentrated in certain areas of the mountainsides, it is these areas that frequently become susceptible to avalanching (Kotlyakov and Plam, 1965). According to McClung and Schaerer (1993), loose snow avalanches are the result of dry snow falling in calm conditions, however, slab avalanches are more likely to occur when wind speeds exceed a threshold of about 7 m/s. Snow particles are transported by the wind via one or more of the following mechanisms; creep, when particles are rolled along the surface; saltation, when particles jump across the snow surface; and suspension involving particle movement in suspended flow above the surface. This final mechanism transports snow particles at a mean horizontal velocity similar to the surrounding wind velocity (Pomeroy and Gray, 1995).

The total amount of snow transported is a function of the wind speed. According to The Swiss Federal Institute for Snow and Avalanche Research (SLF, 2006) wind transported snow occurs with a wind speed greater than 4 m/s for loose snow and greater than about 10 m/s for denser snow. However, the amount of snow drifting attains a maximum with wind speeds between 50 – 80 km/h (approximately 15 – 20 m/s). Beyond this threshold the re-deposition of snow decreases. McClung and Schaerer (1993) suggest this upper bound to be slightly higher at 25 m/s, stating that wind speeds above this can transport snow high above mountain ridges in plumes which result either in loss due to evaporation, or deposition of snow below expected starting zones producing less significant slab formation. The snow transport – wind speed relationship has been outlined by Pomeroy and Gray (1995), who cite several alternate empirical expressions used to estimate the snow transport rate from wind speed data. The different equations derived are presented in figure 2.2(a) with their associated graphical representation in figure 2.2(b). The general trend, as expected, indicates increasing snow transport with higher wind speed.

There are several explanations for the differences between the formulas expressed in figure 2.2(a) including the integration of the mass flux for different heights. Also, assumptions, measurement techniques and snow surface conditions at the locations varied (Pomeroy and Gray, 1995). The four more recent adaptations (Dyunin and Kotlyakov, 1980; Takeuchi, 1980; Tabler *et al*., 1990; Pomeroy *et al*., 1991) are more closely related as seen in figure 2.2(b). Budd *et al*. (1966), however, presumed greater saltation rates than usually measured hence the variation from the other four expressions (Pomeroy and Gray, 1995).



Equation	Height (m)	Reference	Comments
$\log q_T = -1.82 + 0.089u_{10}$	0.001 - 300	Budd *et al.* (1966)	Antarctica
$q_T = 0.000077(u_{10} - 5)^3$	0 - 2	Dyunin & Kotlyakov (1980)	Siberia
$q_T = 0.0002u_1^{2.7}$	0 - 2	Takeuchi (1980)	Japan: old, firm snow
$q_T = 0.0000014u_{10}^{4.2}$	0 - 10	Tabler *et al.* (1990)	Antarctica/ prairie synthesis
$q_T = 0.0000022u_{10}^{4.04}$	0 - 5	Pomeroy *et al.*, (1991)	Canadian Prairies fallow

Figure 2.2(a) (top) Outlining the equations for calculating the transport rate of blowing snow, q_T in kg/s per meter perpendicular to the wind over a specified height range. u is the wind speed in m/s at the height indicated by the subscript in meters (Pomeroy and Gray, 1995).

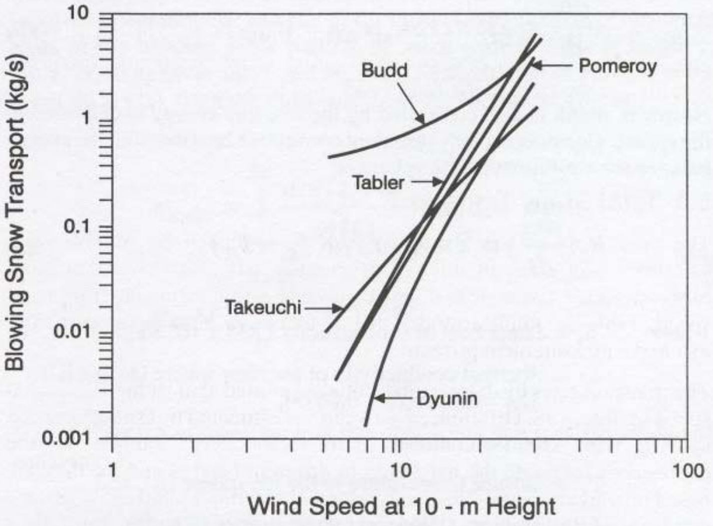


Figure 2.2(b) (bottom) Graphical representation showing total snow transport rate as a function of wind speed at a height of 10 m above the surface. These were calculated from the varying expressions in figure 2.2(a) developed in the locations stated (Pomeroy and Gray, 1995).

According to Barry (1981), the degree of snow transport also varies with such properties as the temperature, size, shape and density of the snow particles and degree of inter-granular bonding of the snow cover. For loose un-bonded snow threshold wind velocity at which the snow is picked up from the surface and transported is approximately 5 m/s (at 10 m), compared with a wind speed of > 25 m/s required to blow dense, bonded snow covers (Barry, 1981).

2.2.2.1 The wind drift factor

When considering the ability to predict avalanches, the application of a wind drift parameter has proved very useful. Both Davis *et al*. (1999) and Hendrikx *et al*. (2005) present this with the use of classification trees. Davis *et al*. (1999) uses the above expressions from Pomeroy and Gray (1995) to derive the wind drift factor as the product of the 24-hour snowfall and wind speed to the fourth power (see equation [Eq. 2.2] below). The 24-hour snowfall is assumed an appropriate index of snow supply to substitute for the constant of unlimited snow supply.

$$\text{wind drift (mm(m/s)}^4\text{)} = \text{precipitation (mm) x (wind speed)}^4 \text{ (m/s)} \qquad \text{[Eq. 2.2]}$$

Davis *et al*. (1999) found that when creating classification trees from a combination of both primary measurements and wind drift data, the two and three day wind drift parameters ranked within the top five factors in every test, often only slightly below the two and three day snowfall and depth parameters. In the study by Hendrikx *et al*. (2005) it was found that creation of classification trees with similar variables as used by Davis *et al*. (1999) showed the first split to be based on the three day temperature dependent wind drift parameter. This indicates the importance of the combined wind drift parameters within classification trees, as Hendrikx *et al*. (2005) concludes that with 78 % accuracy their study correctly classifies avalanche days using only wind speed and a temperature sensitive wind drift parameter.

Kronholm *et al*. (2006b) have also undertaken a preliminary study using wind drift parameters created from an interpolated gridded data set. Only single element trees were created using these combined parameters, however these showed certain wind drift parameters were the best predictor in 12 out of 15 occasions for different avalanche types. In this thesis, similar wind drift parameters will be created with data from the Fonnbu weather station which has not been used previously.

2.2.3 Wind direction

Wind direction is an important factor to observe within close proximity to avalanche starting zones due to variations from local terrain features. In addition it is often necessary to consider the direction and speed of wind in the preceding days to build a fuller picture of snow stability in the area (McClung and Schaerer, 1993). Wind transported snow accumulates in the formation of snow drifts which are strongly influenced by meso- and macro-scale topography. This occurs on the lee side of hills and mountain ridges, in areas of surface roughness and vegetation growth, and also in topographic depressions (Barry, 1981; Pomeroy and Gray, 1995). These features all cause decreases in wind speed and hence saltation and suspension rates. A simplified example is provided in figure 2.3(a) which outlines the process of snow deposition on the lee slope due to deceleration. Alongside this, figure 2.3(b) shows a graphical representation modified by Pomeroy and Gray (1995) of generalised snow deposition with distance from the crest. It indicates the relation to mean mass accumulation on windward and lee slopes with a wind direction from left to right on the figure.

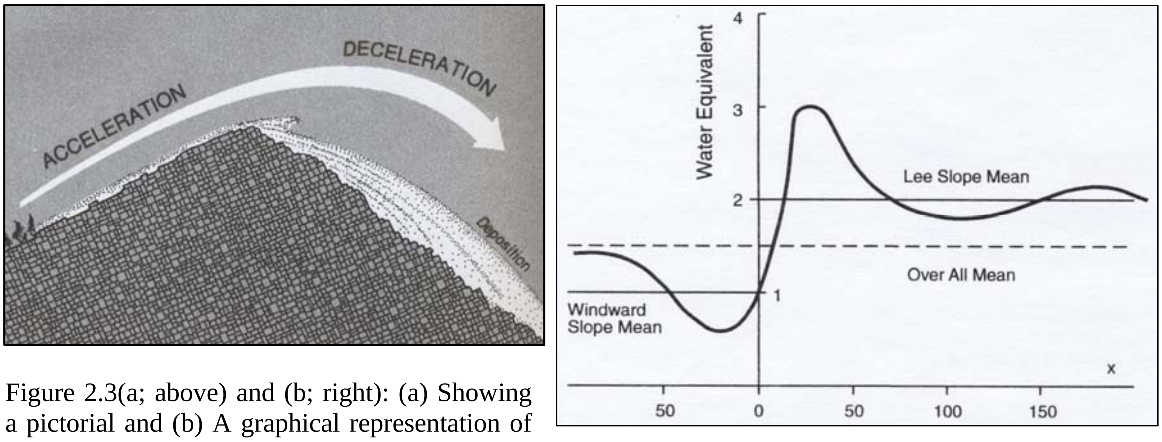


Figure 2.3(a; above) and (b; right): (a) Showing a pictorial and (b) A graphical representation of wind-drift with snow deposition on the lee of the slope as wind direction is from the left to the right of both figures. Source: (a) modified from McClung and Schaerer (1993) and (b) Pomeroy and Gray (1995).

Snow accumulation occurs when snowfall rate is greater than the combination of surface erosion rate and sublimation rate (Pomeroy and Gray, 1995). Furthermore it is stated that the snow accumulation rate is proportional to the fetch distance (Pomeroy *et al.*, 1998). Here it can be noted that fetch distance is not only related to the direction the wind is coming from, but it in turn, has an effect on wind speed as a greater fetch will allow the formation of greater wind speeds.

Bakkehøi (1987) outlines a significant factor which may account for the differing probabilities of avalanche prediction on the five separate paths presented in figure 2.1. This is the relation between prevailing wind direction and aspect of the starting zone. This suggests that the more exposed the starting zone is to snowdrift catchment the more likely the path is to avalanche. To explain further, two avalanche paths with similar inclination and preceding weather conditions may not show a significant correlation in frequency of avalanching as they have differing slope aspects with one accumulating a much higher snowdrift and avalanching more frequently than the other.

2.2.4 Weak layer formation

It is important to mention temperature at this stage as this is the overall controlling factor on what form the precipitation takes. This therefore not only influences the amount of snow-loading during periods of precipitation but also affects the internal structure and stability of the snowpack if water percolation occurs during periods of positive temperatures. Within many mountainous environments daily air temperatures can fluctuate greatly above and below 0˚C, often dependent upon the time of day or night, the amount of cloud cover at these times, altitude and also possible temperature inversions which occur frequently in the mountains. These changes in temperature allow melting to occur, water can then percolate downwards and flow within the snowpack, freeze-thaw cycles can therefore arise during a 24 hour period (McClung and Schaerer, 1993). This melt-freeze metamorphism causes rounded coarse grains to develop with melt water in-between. A weak grain structure is hence formed and this may be the origin of a potential sliding plane for an avalanche. However, if this layer proceeds to freeze before avalanche formation, strong ice crusts can form at the snow surface or in the snowpack, after burial by subsequent snowfall these become possible future sliding surfaces on which weaker layers may slide in an avalanche (SLF, 2006).

Butler (1986) reports on historical avalanches and their meteorological triggers at a study site in Glacier National Park, Montana. The outcome suggested that of 223 avalanches, their triggering factors could be split between four main categories of weather conditions. Of these 223 avalanches, 80 % were a result of changes in temperature, this comprised; 13 % of avalanches caused by a rise in temperature following a heavy snowfall; 27 % of avalanches related to only a rise in temperature above freezing; the majority of avalanches at 40 % related to a rise in temperature coupled with rainfall. To note, the fourth category comprising the remaining 20 % of avalanches are attributed to heavy snowfall only.

2.2.4.1 Surface hoar

This forms due to a relative difference in temperature at the air and snow surface interface. It often occurs at night when a moist oversaturated air mass hangs above a colder snow surface. Condensation of the air mass water vapor produces a layer 1mm to several cm thick of surface hoar "*feathery crystals*" on the snow (McClung and Schaerer, 1993; Barry, 1981). This is a predominant factor in avalanche formation which may occur as this weak hoar layer subsequently becomes buried under new snow.

2.2.4.2 Depth hoar

When snow metamorphism occurs at depth in the snow pack, the resulting formation may be a layer of depth hoar. According to Akitaya (1974) these fragile layers of depth hoar form due to a large temperature gradient within the snowpack, and it is these layers which Akitaya (1974) states are significant in avalanche formation within the areas of Hokkaido and Honshu in Japan. LaChapelle (1962) also recognizes the significance of depth hoar in avalanche formation as the re-crystallization of the layer causes a "*deterioration of the load-bearing capacity*" of the snowpack.

2.2.4.3 Radiation and sensible heat

Sensible heat accounts for the heat transferred between the snowpack and atmosphere, the main process for this is turbulent exchange due to wind eddies (McClung and Schaerer, 1993; Barry, 1981). Foehn winds are an example, particularly prolific in the European Alps, in which a down slope wind causes air temperatures to rise and relative humidity to fall (Barry, 1981). This in turn produces significant warming and hence melting of the snow as it blows over the surface.

Jamieson (2004) considers the role of sensible heat in the formation of melted and refrozen layers which he refers to as temperature crusts, these then often become the bed surface for many slab avalanches. His work, undertaken in the Columbia Mountains of Canada, showed that this process of surface melt-freeze metamorphism is most predominant in March and April. This is due to the requirement of above freezing air temperatures and the necessity for a continuous, well established snow cover. Jamieson (2004) also states that an increase in wind speed associated with the increase of sensible heat exchange between the air and snow surface is considered to produce more surface melting on windward facing slopes.

Radiation interaction with snow cover is stated as having a greater importance in relation to snowmelt situations than sensible heat (Male, 1980) and has been given a concise definition by McClung and Schaerer (1993). This radiation includes short-wave radiation from the sun and long-wave radiation from terrestrial sources particularly the earth itself and clouds. The varying balance between these sources of radiation results in rapid temperature changes at the snow surface. This is significant for the formation of avalanches via the creation and subsequent burial of weak layers caused by surface warming and cooling. Another point to note regarding heating by short-wave radiation is that the percentage absorbed into the snowpack increases by approximately 10 % when the snow surface is wet rather than dry. Also the depth of radiation penetration increases in wet coarse grained dense snow which is a rapid instigator of instability in wet snow (McClung and Schaerer, 1993).

2.2.5 Rainfall

A final meteorological factor to mention is rainfall. This is an important factor to consider, primarily as this adds weight to the snowpack (McClung and Schaerer, 1993). Also, Ambach and Howorka (1965) have stated that initiation of wet snow avalanche activity occurs when the liquid water content of the snow reaches 7.5 % by volume, however this must also be associated with "*large positive values of the heat budget*".

Another effect of rainfall on snow cover is, if the snow is relatively warm, water percolates downwards having the effect of changing mechanical properties of the subsurface layers. The water may refreeze deep within the snowpack causing ice lenses to form, these lend themselves to ultimately being a surface on which weaker layers above can slide. If the snow surface is cold the rain may refreeze immediately on the surface and, on burial by subsequent snow, this layer may have the same effect as stated previously, possessing properties of a subsequent avalanche sliding surface (McClung and Schaerer, 1993).

Jamieson (2004) concurs with the above statements and refers to the snow surface, which becomes wet due to rainfall and subsequently refreezes, as a rain crust. Additionally the effect of wind during a rain storm is mentioned, in which wetter and often thicker layers result on windward slopes than the lee sides as more rain is received per unit area. This may therefore hold the characteristics of being a more continuous future avalanche sliding surface layer.

2.3 Influence of climatic factors for avalanche formation

The preceding review in section 2.2 has detailed the effect that short term weather phenomena and patterns have on influencing avalanche formation. However, it is necessary to introduce climate induced factors which may have a bearing on frequency of avalanche occurrence during certain years. Climatic cycles are important to note as these may have a longer lasting influence over the weather of particular regions, rather than the annual cycle normally associated with certain weather patterns. Changes in climate have therefore been stated to have an effect on avalanche return periods by Keylock (2003) and Lied *et al*. (1998). The North Atlantic Oscillation (NAO) is just one such example. It is defined as the standardised sea-level air pressure difference between the Azores high and Icelandic low (Met office, 2006; Keylock, 2003), and is considered an important factor for climate fluctuations in the Northern Hemisphere. A positive NAO refers to strong westerly winds and warmer and wetter than average conditions for northern Europe. The opposite occurs for a negative NAO, that is, dryer, colder and less windy conditions in northern Europe.

There is some discussion regarding the influence of the North Atlantic Oscillation on weather and climatic factors, and hence the consequence for avalanche formation. In Norway a correlation has been suggested between the NAO and temperature and precipitation events (Hanssen-Bauer and Førland, 2000). Keylock (2003) also discusses the influence of the NAO on avalanche prediction and formation in Iceland. He states that the recent positive correlation between the NAO index and monthly precipitation in Iceland is also observed in the glacial mass balance records of Scandinavia. On this vein of thought "*an avalanche release is not a simple function of snowfall, there would appear to be a possibility that avalanche activity is also correlated with the NAO*" (Keylock, 2003).

Although the above paragraphs signify the NAO as the key role in the warmer and wetter weather experienced over northern Europe, it must be acknowledged that intensification in the anthropogenic greenhouse warming effect may have similar consequences on the climate. Ulbrich and Christoph (1999) highlight this with various climate models to indicate that the storm track bringing wetter weather across northern Europe intensifies due to this effect.

2.4 Norwegian weather and climate

In general western Norway is characterized by a maritime climate with relatively mild temperatures and high precipitation. Most of this precipitation occurs in late autumn, winter and early spring. The Norwegian Meteorological Institute (2006) state that there are three different categories into which Norwegian precipitation events can be subdivided, these are frontal, orographic and showery precipitation. In general, frontal precipitation accounts for the majority of rainfall across Norway forming at the polar front, this is where the moist and warmer air from the south meets colder dryer air from the north. Vertical currents in unstable air provide showery precipitation, however this effect is most dominant in the inner areas of Norway including Østlandet and Finnmark and occurs mostly during the summer months. At times the showery precipitation can coincide with and hence intensify orographic or frontal precipitation.

In western Norway the predominant form of precipitation is orographic, created when warm air currents are forced to rise over the mountains as they move inland, this causes cooling and subsequent condensation and precipitation. According to the Norwegian Meteorological Institute (2006) this effect gives more precipitation than would otherwise be expected, and provides maximum effect 50 km inland. In terms of numerical values average annual precipitation of 3575 mm is measured at Brekke in Sogn and Fjordane County.

Winter months see heavy snowfall and fluctuating temperatures, the air temperature often rising above 0˚C causing periods of high snow accumulation to be interrupted by high intensity rainfall (Blikra and Nemec, 2000). As summarised by McClung and Schaerer (1993) new snowfall closely followed by rainfall can cause major avalanche events. In short, these varying weather conditions lead to fluctuating instability of the mountain snowpack and frequent avalanche events. McClung and Schaerer (1993) go on to outline that avalanches often closely follow the winter storms with failure occurring near the surface of the new snow. It is this direct-action avalanching, which occurs during or shortly after winter storms that will form the bulk of the investigation to further this thesis report. The predominant focus will therefore be on preceding snow accumulation and parameters of wind speed and direction.

2.5 Avalanche forecasting techniques

McClung and Schaerer (1993) give a detailed approach on how to forecast and ultimately predict avalanches. This involves the use of both numerical and descriptive data the former being more measurable and more user friendly for further analysis.

Meteorological data is of this kind, with real time observations and an almost continuous data flow available with the use of data loggers and computers. Forecasting on different scales has been given some discussion as this can depend upon accessibility and practicality of a forecast over a certain area. It is stated that regional forecasts over entire mountain ranges rely heavily upon meteorological data being principally office based, whereas local forecasting applies to "*avalanche prediction on a smaller scale, usually for an area of less than 100 km*^{2}" (McClung and Schaerer, 1993).

There are a number of avalanche prediction methods in use with varying degrees of accuracy and suitability. Conventional avalanche forecasting uses a vast array of data and information but generally without the use of numerical and analytical procedures, instead, results rely on "*intuition, experience and local knowledge*" (McClung and Schaerer, 1993). More recently, success has been achieved with the application of numerical avalanche prediction methods for local forecasting, two such examples are discriminant analysis and nearest neighbours. Bois *et al*. (1974) achieved good indicative results using discriminant analysis to distinguish between dry snow avalanche days, wet snow avalanche days and no avalanche days. The technique of discriminant analysis involves firstly finding the variables which best discriminate between groups and secondly classifying the given events by assigning them to the different groups. Bois *et al*. (1974) outlines a number of advantages to using this technique, including the fact that more than one subset can exist therefore both wet and dry avalanche days can be distinguished in addition to no avalanche days. Also there is the possibility to control and reduce the number of days with no avalanches, as otherwise this large group would dominate the statistical analysis.

Nearest neighbour models use an archive of historical measurements in order to find the best match to current measurements for use in avalanche prediction. Brabec and Meister (2001) and Gassner and Brabec (2002) outline examples in Switzerland, stating that on a local scale the nearest neighbour technique provides the ten most similar days for a given situation. The present avalanche danger is then indicated by the ten historically observed avalanches. Avalanche forecasting is also practised on a regional scale in Switzerland using the nearest neighbour technique. After cross-validation, results show that although only 52 % of the days are in agreement with conventional estimates of hazard levels, 96 % were within one hazard level.

2.5.1 Classification trees

A final statistical procedure to be discussed is that of classification trees. These are a good aid for outcome prediction and can also provide patterns and description to the underlying structure of certain data sets (Davies *et al*., 1999). In terms of their forecasting capabilities, classification trees are used to predict a dependent variable, in this case avalanche days or non avalanche days from a number of predictor variables i.e. preceding weather parameters (Hendrikx *et al*., 2005; StatSoft, 2003). As described by Davis *et al*. (1999) a classification tree is constructed by an algorithm "*recursively partitioning the data into increasingly homogenous subsets until each subset contains a small number of cases*". Pruning then selectively recombines some branches depending on their similarity in-order to achieve the desired criteria of levels and complexity. The results depend rather on historical accuracy of the data than on a pre-determined confidence level (Hendrikx *et al*., 2005).

Although there are many more statistical procedures which could be mentioned, with regards to the scope of this thesis and available related literature, the most important methods have been outlined above for reference and further use at a later stage. It is recognised that there is a lot more to the classification tree method than has been included in this initial section. Also as this is perhaps the most appropriate technique to implement with regards to avalanche day prediction based on weather data, more detail for this method will be provided in the later sections of chapters 3 and 4.

Chapter 3: Methodology

3.1 Study area

The mountainous topography and climatic conditions of western Norway create an ideal setting to undertake avalanche research. The area to be focused on in this thesis surrounds Strynefjell in Sogn and Fjordane. Figure 3.1 shows the location of the study area at Strynefjell in relation to the outlying district of Stryn in Sogn and Fjordane.

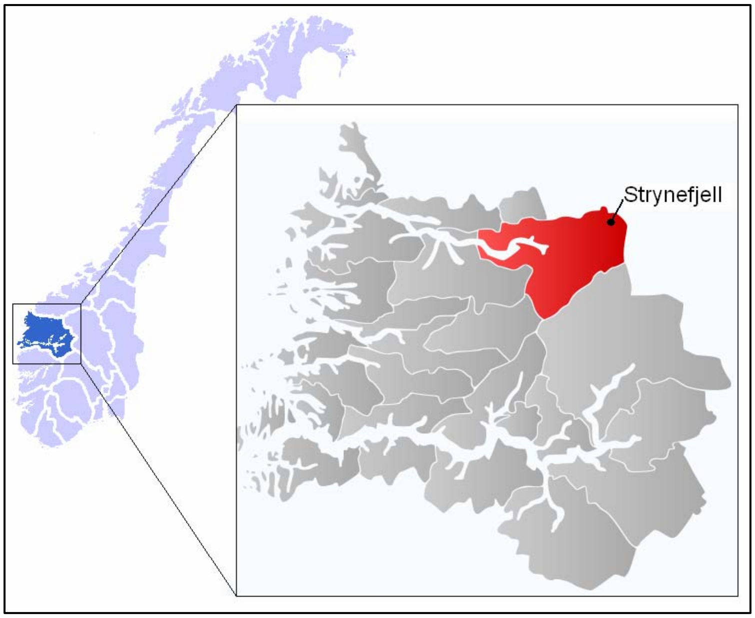


Figure 3.1: Strynefjell indicated within Stryn kommune (red) within Sogn and Fjordane fylke (blue) in western Norway. Map adapted from Wikipedia (2007a).

This site has been chosen due to the availability of data from the Norwegian Geotechnical Institute (NGI) with regards to weather records and related avalanche event information, as their avalanche research station has been based at the head of Grasdalen since 1973. This U-shaped valley is above the tree-line and is therefore sparsely vegetated with small hardy shrubs as would be expected at this altitude on frequently avalanching terrain. The maritime climate of the region is influenced largely by cyclonic activity in the Atlantic Ocean (Bakkehøi, 1987). The mean value of precipitation during the winter period (November – April) has been stated as 855 mm, giving maximum snow depths recorded between 1.3 m and 4.7 m (Bakkehøi, 1987).

The main route through this area is the RV15 which connects the western coastline to the inland towns of Lom and Otta to the east. This road is threatened by numerous avalanches as it follows the valley. Several tunnels, three of which are within the area surrounding Grasdalen, are used to bypass peaks of up to 1700 m. Figure 3.2 shows a detailed map of the study area with avalanche release zones depicted with a red dot and numbered from 1 to 51. Each of these 51 different paths have individual names (a list of which is provided in Appendix 1), however for map clarity only the numbers have been indicated at this stage. These numbers also indicate a ranking with avalanche path 51 releasing most frequently and avalanche 1 being the least frequent release path.

To illustrate the topography and landscape of the area, figure 3.3 has been included. This photograph has been taken from near the NGI research station (located by avalanche 1 in figure 3.2) and looking south down the valley. A 100 m long gallery over the road can be seen which protects the road from the most frequent avalanche path 51 (Sætreskarsfjellet) which is the slope to the right in the picture. The peak in the centre of the photograph is Raudnova, at 1665m and plays host to avalanches 48 (Svartefjellet), 41 (Raudnova NW), 33 (Raudnova top NE), 24 (Raudnova top N) and 19 (Raudnova W). On the left of the valley in figure 3.3, the main avalanche paths are 37 (Fonnbu NE for) and 35 (Svartebardskaret).

Although long stretches of the road follow tunnels and the strategically placed gallery, there is a 1.5 km section towards the head of Grasdalen which is predominantly uncovered. Part of this section before the gallery can be seen in figure 3.3; this is exposed to avalanches along the whole length (apart from the 100 m gallery section). The road can therefore be blocked two or three times during the winter season, however closure of the road is often accomplished before avalanches block it due to the current forecasting measures in place (Bakkehøi, 1987).

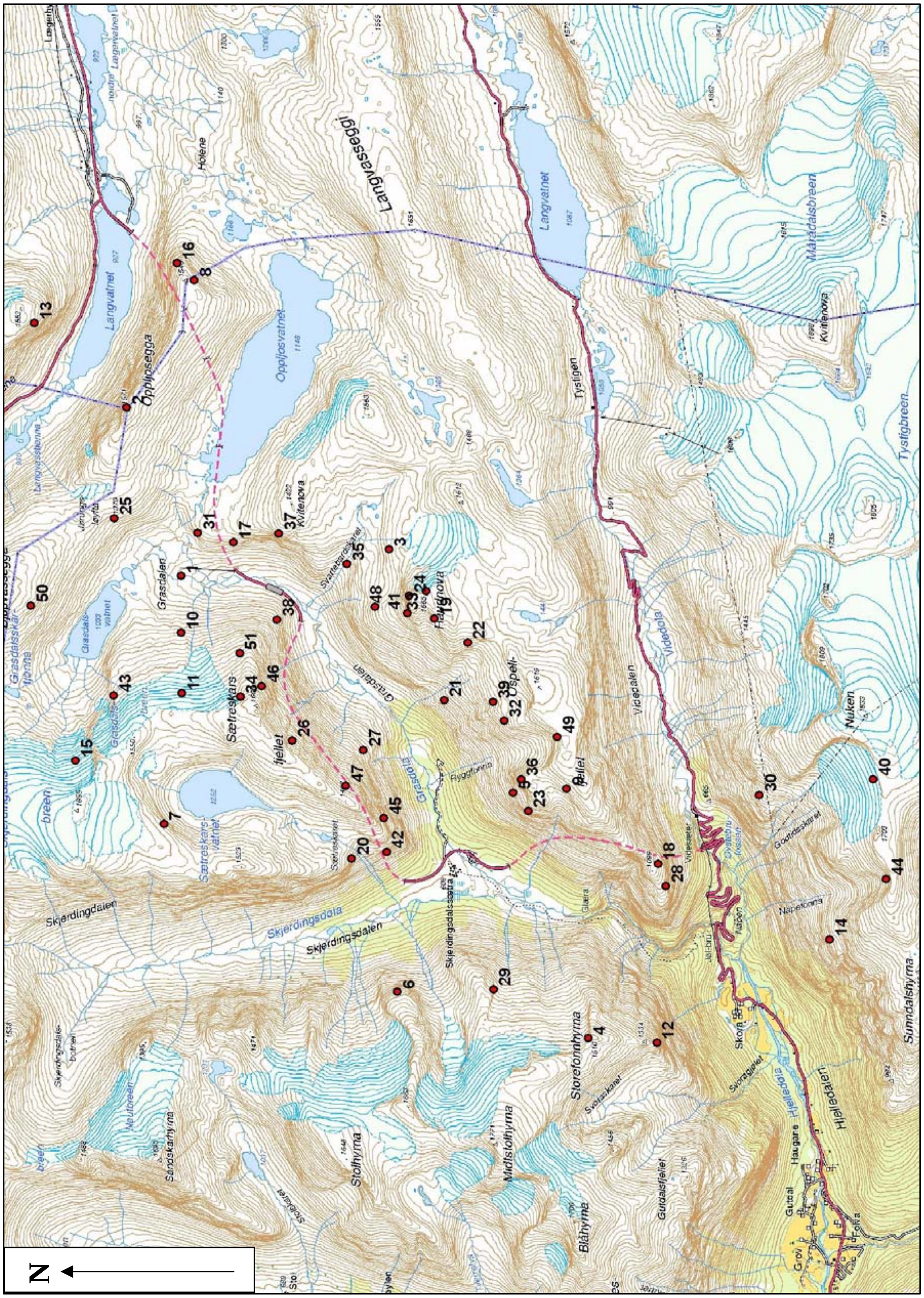


Figure 3.2: Grasdalen study area with avalanche release zones marked in red and numbered 1 (least frequent) to 51 (most frequent). Map scale 1:35,000.

Figure 3.3: Grasdalen from the NGI research station looking south, the RV15 can be seen in the valley bottom and Raudnova peak in the background.

3.2 Data

Due to the long time series required and amount of data necessary for analysis, this was not collected personally. Instead, data sets were provided from a variety of sources including the Norwegian Geotechnical Institute (NGI), the Norwegian Meteorological Institute, and the Bjerknes Centre for Climate Research in Bergen which has provided data from the Environmental Modelling Centre (NCEP). These raw data sets are listed below:

- Avalanche data (*Grasdalen_normalskred_all_data_NGI*)
- Gridded temperature and precipitation data (*Grasdalen_Klimadata_skreddager_METNO*)
- Gridded wind data (*Grasdalen_vinddata_skreddager_BCCR*)
- Fonnbu weather station data (*Avalmet_alldays_NGI*)

3.2.1 Avalanche data

The avalanche data has been collected via direct observations by NGI personnel within the Grasdalen area, this is generally done during or just after the storm periods. The data set includes information regarding location, path, release date/time, triggering mechanism i.e. natural or artificial, and type i.e. wet or dry avalanche. The date and time of each avalanche occurrence is estimated by checking the new snow depth on the avalanche debris if possible, and accuracy in terms of error bands are stated for the assigned avalanche date (Bakkehøi, 1987).

The raw data set consists of 1048 avalanche occurrences within the period 1974 to 2002, however during this time data collection and recording practices have changed (Kronholm *et al.*, 2006a), and as can be expected on such a large data set there will be some errors and missing values within the data. The total number of avalanche occurrences has been reduced to 805 occurrences for analysis, this is due in part to errors and missing values. However, more significant reasons for this were that avalanches with a date accuracy of greater than +/-12 hours were considered too inaccurate to correlate with daily weather observations. In addition, for obvious reasons only natural release avalanches are considered as artificially detonated avalanches will produce inconsistencies with regards to avalanche prediction due to weather phenomena. Finally, in relation to avalanche type, preliminary analysis was undertaken on the 805 events with the inclusion of both wet and dry avalanches. However, as this thesis will focus on the topic of the wind drift parameter which is considered more important for dry avalanche triggering, all wet avalanches are disregarded at the later stages of analysis. It must be noted that there are however a number of avalanches of unknown type, these are to be left in the data set, as excluding them would leave too little data for worthwhile statistical analysis.

3.2.2 Gridded temperature and precipitation data

This weather data has been provided by the Norwegian Meteorological Institute and consists of mean air temperature and precipitation on the day of the avalanche and, additionally, the three and five day sum of precipitation with the final day ending on the day of the avalanche. These values are the interpolated results for the Grasdalen area from a nationwide 1 km by 1 km grid. The weather station locations from which this data was generated are located irregularly across Norway, with a less dense network at higher elevations. Unfortunately, this produces distorted values for precipitation particularly at higher altitudes (Norwegian Meteorological Institute, 2004). In order to use this data set, the data has been cross referenced to correlate for each of the 805 avalanche occurrences to be analysed between 1974 and 2000.

3.2.3 Gridded wind data

The wind data is provided by the Environmental Modelling Centre (NCEP) and has been similarly interpolated into gridded results as outlined for the climate data above, providing a grid of 250 km by 250 km (Kistler, 2001 in Kronholm *et al.*, 2006b). This data set has been modified to provide values of average wind speed and maximum wind speed on the day of the avalanche and also across the three and five days prior to avalanche events with the final day being that of the avalanche day. In addition, mean wind direction and maximum wind direction is given on the day of the avalanche, plus likewise as above, across the three and five days prior to avalanching. The measurements apply to a height 10 m above the ground surface, and locations picked from the grid correspond with the highest point in the release area of each avalanche path (Kronholm *et al.*, 2006a). As with the climate data, this wind data was also sorted to correspond with the 805 avalanche events.

3.2.4 Fonnbu weather station data

This data combines actual recorded values of weather parameters on a daily basis from the NGI weather station at Fonnbu, Grasdalen which can be seen pictured in figure 3.4. This is a valley site and hence is likely to experience wind tunnelling which may result in distortion of some data. The data collected here includes daily precipitation which has been used to determine the one, two, three, four and five day sum of precipitation over the preceding days. Temperature is also recorded to provide data for the mean, maximum and minimum daily values. Mean and maximum wind speed are also measured from here at a height 10 m above the ground surface (however this height varies with snow depth). Wind direction data has not been provided with this data set as at the time of writing it is unavailable. It is also believed to be of little importance as the valley location is likely to channel the wind resulting in alteration of the true wind direction to one of two main directions.

The Fonnbu data set also includes recording of various snow depth parameters over the preceding days. Avalanche days are also highlighted to state if an avalanche occurred on that day and if so how many wet avalanches, dry avalanches, or avalanches of unknown type. This data is a full time series from 1974 to 2000, and has been collected both manually and by automated equipment. Only data from the winter seasons was selected for further analysis, for ease this was defined as from 1st December to 30th April. Over this period, this resulted in a total of 314 days recorded as avalanche days and the remaining 3588 days recorded as non avalanche days.

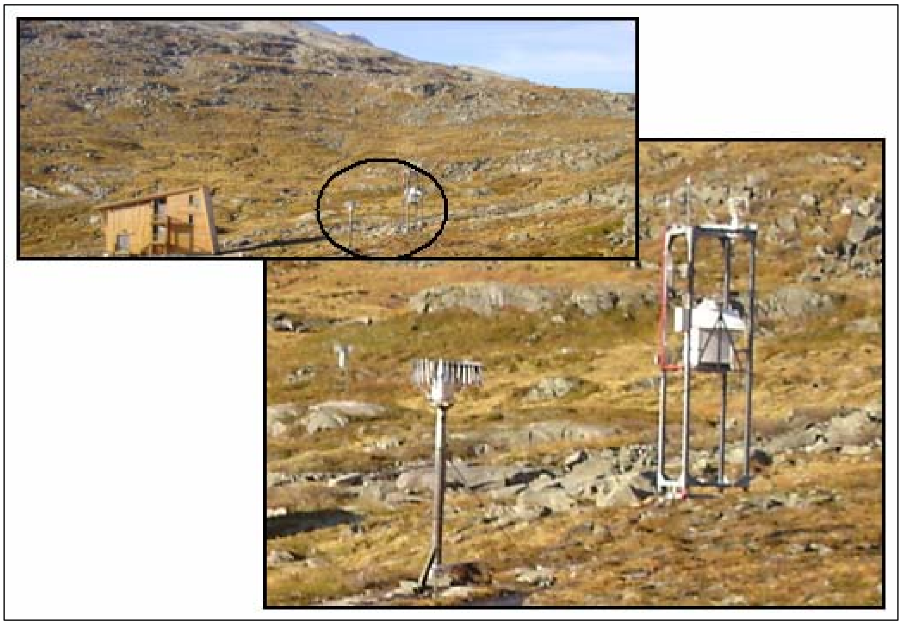


Figure 3.4: Fonnbu NGI avalanche research station at the head of Grasdalen, showing the location of the weather monitoring station circled and in the foreground. The precipitation gauge with wind shield is seen to the left. The anemometer measuring wind speed is placed 10 m from the ground on the structure to the right in the figure.

3.3 Data errors and discrepancies

With regards to the measured data values from the weather stations it must be mentioned that certain problems arise due to environmental factors. One of the most prevalent is that of riming which is the process describing the freezing of vapour and crystals on to the anemometer measuring equipment leading to inaccuracy and ultimately an inability to record the present wind speed conditions. McClung and Schaerer (1993), however, state that the use of radiant heaters can keep the equipment free of rime. In addition to this, uncertainties may occur with regards to precipitation measurements during strong wind. As can be seen in figure 3.4 there is a wind shield around the precipitation gauge which will help to some extent, however snowfall is more prone to the effects of wind than rainfall so inaccuracies may be acquired dependent upon the wind conditions.

Another problem associated with the data sets is the overall accuracy which varies depending on such things as qualitative observations by different personnel under differing conditions. The method of data recording has obviously changed over the 25-year period, as have the personnel and recording equipment. The frequency of taking measurements may also have changed over this time and as a consequence the frequency of checking that the equipment is in working order and recording accurately may have reduced, particularly as over time the equipment has become more automated. So for example; if problems arise with certain equipment with it being automated and not manually measured on a daily basis, it may be several weeks before any personnel visit the site and are able to rectify the problem.

An additional problem related to the recording of data by personnel in the study area is that of subjective results. An example of this being in the recording of avalanche type, this may often have been done by way of visual observation from almost 2 km away. From this distance it can obviously be problematic to state if the avalanche is either wet or dry.

With such large data sets as those to be used in this thesis there is bound to be a certain degree of error or discrepancy within and between the data sets. This may be down to initial input error or the introduction of some errors during processing of the data. It is however presumed that the overall processing of the data for quality improvement has outweighed the impact of errors from erroneous values that may have been introduced inadvertently. Finally, cross referencing of data was done between the climate, wind and avalanche data by way of an avalanche ID number, meaning each of the 805 slide events, after reduction of unsuitable data had an individual number (between 1 and 1018). In the case of the combined data from the Fonnbu weather station, avalanche events did not have an individual ID, just a total number of events each day, it was therefore not so easy to cross reference between this and the other data sets and some discrepancies in further analysis are bound to arise.

3.4 Statistical methods and techniques

3.4.1 Exploratory data analysis

Exploratory data analysis is first undertaken for each data set in order to obtain a clearer understanding of the data involved. This has included a number of basic histograms and scatter plots of each weather parameter over the 25 year period. In addition to this the parameters within and between data sets were compared to one another to see what correlation exists if any. This was carried out by plotting some of the parameters against each other, and also by comparing statistical moments of mean and standard deviation, in addition to minimum and maximum values, to indicate the degree of spread of the data. Once individual parameters within the data sets have been analysed, the first three data sets i.e. the avalanche data (3.2.1), and the two sets of gridded data; temperature and precipitation data (3.2.2) and wind data (3.2.3) were combined to provide one large set of gridded data containing all weather parameters.

3.4.1.1 Calculation of wind drift data

Once the gridded data was combined into one large data set it was possible to create certain wind drift parameters using the expression derived by Davis *et al*. (1999) and based on the work done by Pomeroy and Gray (1995). This is defined as the product of precipitation and wind speed to the fourth power (see [Eq. 2.2]) and was done over the one day, three day and five day periods for the gridded data. This was also carried out with the addition of the preceding day precipitation and two and four day sum for the Fonnbu data set. Tables 3.1(a) and (b) provide a summary of these parameter codes and calculation descriptions for each of the data sets.

Gridded Data		
parameter code	***Description***	***Units***
rrwndspd1	rr1day x $(\text{wndspd1day})^4$	$\text{mm}(\text{m/s})^4$
rrwndspd3	rr3dayx $(\text{wndspd3day})^4$	$\text{mm}(\text{m/s})^4$
rrwndspd5	rr5day x $(\text{wndspd5day})^4$	$\text{mm}(\text{m/s})^4$
rrwndspdmax1	rr1day x $(\text{wndspdmax1day})^4$	$\text{mm}(\text{m/s})^4$
rrwndspdmax3	rr3day x $(\text{wndspdmax3day})^4$	$\text{mm}(\text{m/s})^4$
rrwndspdmax5	rr5day x $(\text{wndspdmax5day})^4$	$\text{mm}(\text{m/s})^4$

Table 3.1(a): Showing the parameter codes assigned for the combined wind drift data and a description of their calculation for the Gridded data set.

Fonnbu Data		
parameter code	***Description***	***Units***
RRFFM0	RR0day x $(\text{FFM0day})^4$	$\text{mm}(\text{m/s})^4$
RRFFM1	RR1day x $(\text{FFM1day})^4$	$\text{mm}(\text{m/s})^4$
RRFFX0	RR0day x $(\text{FFX0day})^4$	$\text{mm}(\text{m/s})^4$
RRFFX1	RR1day x $(\text{FFX1day})^4$	$\text{mm}(\text{m/s})^4$
RRFXM0	RR0day x $(\text{FXM0day})^4$	$\text{mm}(\text{m/s})^4$
RRFXM1	RR1day x $(\text{FXM1day})^4$	$\text{mm}(\text{m/s})^4$
RRFXX0	RR0day x $(\text{FXX0day})^4$	$\text{mm}(\text{m/s})^4$
RRFXX1	RR1day x $(\text{FXX1day})^4$	$\text{mm}(\text{m/s})^4$
RRFXXmax1	RR1day x $(\text{FXXmax1day})^4$	$\text{mm}(\text{m/s})^4$
RRFXXmax2	RR2day x $(\text{FXXmax2day})^4$	$\text{mm}(\text{m/s})^4$
RRFXXmax3	RR3day x $(\text{FXXmax3day})^4$	$\text{mm}(\text{m/s})^4$
RRFXXmax4	RR4day x $(\text{FXXmax4day})^4$	$\text{mm}(\text{m/s})^4$
RRFXXmax5	RR5day x $(\text{FXXmax5day})^4$	$\text{mm}(\text{m/s})^4$

Table 3.1(b): Showing the parameter codes assigned for the combined wind drift data and a description of their calculation for the Fonnbu data set.

3.4.2 Cumulative probability plots

In line with probability plots created by Bakkehøi (1987), similar plots were created for the one, three and five days sums of precipitation for both the gridded and Fonnbu data, along with a selection of probability plots using the combined wind drift factors. The probabilities for these plots were computed by summing individual events within a selected range of values and calculating this as a percentage of the total number of events plus one. This can be outlined simply in the equation below [Eq. 3.1].

$$P = \frac{\Sigma n}{n+1} \times 100 \qquad \text{[Eq. 3.1]}$$

Probability plots were created for the ten most frequently occurring avalanche paths i.e. numbers 51 to 42 (inclusive) from the gridded data set (the list of corresponding avalanche names is given in Appendix 1). These top ten were chosen not only because they are the most frequently occurring, but they also provide a good representation of the range of paths in existence in this area and are spread along the valley with differing aspects. These ten plots were depicted alongside corresponding Fonnbu data. All of these plots used data from dry avalanche days only as this was the case for Bakkehøi (1987).

3.4.3 The Kruskal-Wallis test

Once analysis of both data sets had been undertaken in the manner above, statistical procedures were then carried out involving the ability to discriminate between avalanche days and non avalanche days. From this stage onwards only the Fonnbu data set was used, as this contained both avalanche and non avalanche day data, and has also not been used previously for further analysis.

The Kruskal-Wallis (K-W) test was used to look at each parameter in the Fonnbu data set individually. This test was used as it is a non parametric test for two or more samples (R Development Core Team, 2005; Mathworks, 2007). The null hypothesis being that there is no significant difference between data values on avalanche days and non avalanche days. This was accepted if the probability exceeded the 5 % significance level ($p > 0.05$). The alternative hypothesis states that differences between avalanche day data and non avalanche day data are greater than can be expected from random variation.

As the K-W test is being used for exploratory purposes only, avalanche days were classed purely if there was an avalanche, independent of the type i.e. wet, dry or unknown. Also, of the 314 avalanche days in the Fonnbu data set, 294 of these were used, the remaining 20 were discounted as these were during the time period 1984 to 1988 when data collection was at its most inconsistent with numerous unknown values across the weather parameters. The number of non avalanche days had to be reduced from 3588 to 294 so that the data sets were equal in length. Non avalanche days were picked along the time series at random, between avalanche days. If avalanche days had been further split into wet and dry avalanche days, only 68 wet avalanche days are recorded over the time period meaning 3588 non avalanche days would need to be reduced even further. It was felt that this would restrict the data too much at this stage of the analysis.

3.4.4 Classification trees

3.4.4.1 Background theory

Since the K-W test only looks at individual parameters within the data set, this may be considered too simplified an approach to deal with all the variations which contribute to avalanche occurrences. Classification trees are therefore a valuable tool for multivariate analysis which allow the incorporation of all parameters with the most significant parameters ranked at the top of the tree. There are a number of advantages of classification trees as a method of data mining, these include their ability to handle both discrete and continuous variables, and also their flexibility in handling data sets with missing values (Weka, 2007). In addition, no assumptions are made regarding the distribution of the underlying data, classification trees are therefore categorised as non-parametric procedures. According to Hendrikx *et al.* (2005) they can work on smaller data sets than those required for nearest neighbour analysis, also, what seem complex interactions can become clearly interpretable by pictorial representation. The transparent structure of classification trees can reveal relationships showing how one variable depends on another, this is highlighted if a variable appears more than once in different parts of the tree. Clear interpretation is also possible due, in part, to the fact that at each non-terminal node decisions are based on just one predictor variable (Weka, 2007). This method of data analysis is also very effective where relationships between data are hierarchical or non-linear, in addition, the ability to over-fit a model can highlight data that may otherwise be overlooked. However, it is also possible to prune the tree to a "*statistically defensible size*" by means of cross-validation and cost-complexity (Davis *et al*., 1999). A final advantage to be mentioned is the robust nature of the model to outliers which become isolated in terminal nodes and so do not affect subsequent splitting of the tree (Hendrikx *et al*., 2005; Weka, 2007).

Although the above provides an exhaustive selection of advantages to this method, there are several disadvantages which must be considered at this point. One factor to consider is that classification trees can often be described as unstable due to their production of very heterogeneous results, this can occur due to slight variations made to their growing methods. Also small variations in the data sets applied are particularly important to note when using randomised selections of data; one tree can look very different to another tree by just using a different randomised sample of data (Weka, 2007). This is especially true when there is some overlap across the classified groups with some values lying close to each other. Some difficulty may arise with interpretation of the tree if it becomes too large and complex, this often depends on the number of splits performed, too many splits result in an overly complex and possibly unrealistic model, however it is necessary not to simplify the tree by restraining the splitting too much (StatSoft, 2003). Another disadvantage to note is stated by Weka (2007) that although the models may excel when applied to classification problems, they may not be so accurate in relation to estimation tasks and analysis. Finally classification trees are "*computationally expensive to train*" (Weka, 2007), meaning that a larger data set requires a much greater number of operations in order to grow a tree than a smaller data set.

3.4.4.2 Practical application

As it is necessary to split avalanche day data and non avalanche day data in the classification tree procedure, the Fonnbu data set was therefore used as this consisted of meteorological data on all days during the winter season. The R language and environment software (R Development Core Team, 2005) was used to create the classification trees. This required running the script outlined in Appendix 3 which has been modified to apply to the needs of this data set. It was decided to only look at the split of the classification trees in the case of dry avalanche days versus non avalanche days. At this stage of the investigation wet avalanche days were excluded as the wind drift parameter was to be included which is considered to have more bearing on dry avalanche formation than wet avalanche formation. The parameters used for tree creation are outlined in table 3.2 which includes both the observed data and also the combined wind drift parameters.

Precipitation	Wind	Combined wind drift	Temperature	Snow depth
RR0day	FFM0day	RRFFM0	TAM0day	SS0day
RR1day	FFM1day	RRFFM1	TAM1day	SS1day
RR2day	FFX0day	RRFFX0	TAX0day	SSdif1day
RR3day	FFX1day	RRFFX1	TAX1day	SSdif2day
RR4day	FXM0day	RRFXM0	TAN0day	SSdif3day
RR5day	FXM1day	RRFXM1	TAN1day	SSdif4day
	FXX0day	RRFXX0		SSdif5day
	FXX1day	RRFXX1		
	FXXmax1day	RRFXXmax1		
	FXXmax2day	RRFXXmax2		
	FXXmax3day	RRFXXmax3		
	FXXmax4day	RRFXXmax4		
	FXXmax5day	RRFXXmax5		

Table 3.2: Showing all the parameters submitted for selection by the classification trees. A full description of each can be seen in the glossary of terms.

By running the script (Appendix 3) firstly all the data was loaded, and from this a number of dry avalanche days were selected dependent upon these days having less than 20 unknown meteorological values and one or more dry avalanches during the 24 hour period. A similar procedure was carried out for the non avalanche days, however, with each day having no unknown meteorological values. This reduced the non avalanche days to 742 days, a random selection of these days were then needed to equal the number of dry avalanche days at 143. A binary classification tree was then built with the ‘rpart package’ in R (Therneau and Atkinson, 2005) which uses all 286 observation days; 143 dry avalanche days and 143 non avalanche days.

The tree was grown using Gini values to decide parameter splitting and threshold values. The Gini index is a measure of impurity for a given node and is at a maximum when data is evenly distributed among the classes, and becomes zero when only one class exists at a node (Breiman *et al*., 1993 in Hendrikx *et al*., 2005). In simple terms the Gini rule selects a single group of as large a size as possible, subsequent nodes are then segregated in the same manner until further divisions are not possible (Wikipedia, 2007b). Node heterogeneity was measured by deviance, in the case for this thesis it is a default measure which is used to stop the tree growth. Additionally, misclassification of dry avalanche and non avalanche days were weighted equally. To explain this concept further, on some occasions more accurate classification may be required for some classes rather than others, purely as an example, avalanche days may need to be more accurately predicted than non avalanche days i.e. there is less impact in avoiding a non avalanche day in comparison to the greater impact of not avoiding an avalanche day. In this case greater misclassification weighting could be applied for misclassifying an avalanche day as a non avalanche day than for misclassifying a non avalanche day as an avalanche day (StatSoft, 2003). Equal weighting, being less complex, was however applied for classification trees in this thesis as more concern was placed on the splitting parameters themselves rather than prediction accuracy and misclassification.

The trees were initially grown to their full extent to give perfect classification, however this degree of accuracy often produces a tree too complex to interpret particularly for prediction purposes. They were therefore pruned to an optimal size which is more manageable and meaningful using 10-fold cross-validation and a complexity parameter of 0.03 which seemed the most suitable for the size and accuracy of the trees required. According to Hendrikx *et al*. (2005) this pruned tree is a descendant of the original tree, however makes a summary with the splits that provide maximum correct classification and minimum misclassification. From the pruned tree it was then possible to find out the number of misclassified events at each node, that is to say; the number of dry avalanche days wrongly classed as non avalanche days and vice versa. This misclassification can be added up across the whole tree to produce an overall measure of accuracy.

The script has been run 50 times therefore using 50 random selections of non avalanche days out of the 742 available from the Fonnbu data set, which has not been done previously. This was carried out in order to see how stable the classification trees are with varying data values. For each of the 50 pruned trees the parameters used for the first, second and third splits were recorded along with their threshold values. These first three parameters are considered the most important splits (Kronholm *et al*., 2006b). Additionally it was noted how many pruned trees contain any of the combined wind drift parameters. Finally, the overall accuracy of each pruned tree is recorded.

The results of these trees were then analysed (see chapter 4). This made it possible to see which parameters showed greatest importance by occurring most frequently within the top splitting positions. Threshold values were also looked at in detail to see if any common patterns of values existed.

Chapter 4: Results

To highlight the data sets used and analysis undertaken on each, table 4.1 is included for reference.

Data Sets	Subsets of Data	Statistical Procedures
Gridded data	All avalanche types	exploratory data analysis
	Dry avalanches only	exploratory data analysis probability plots (rr1,3,5day) probability plots (combined wind drift parameter)
Fonnbu data	All avalanche types (avalanche days only)	exploratory data analysis
	Dry avalanches only (avalanche days only)	exploratory data analysis probability plots (RR1,3,5day) probability plots (combined wind drift parameter)
	Avalanche days and non avalanche days	Kruskal-Wallis test
	Dry avalanche days and non avalanche days	classification trees

Table 4.1: summary table of data sets and corresponding statistical procedures

4.1 Exploratory data analysis

4.1.1 All avalanche types

Initial investigation was carried out on all avalanche types in both data sets. As previously mentioned in section 3.1, of the 805 avalanche events during the 25-year period, these occur on one of 51 known paths. Path number 51 known as Sætreskaresfjellet is the most frequently avalanching path with 90 avalanches between 1974 and 2000 following this route. Figure 4.1 provides the name of all 51 avalanche paths against their frequency. The top ten of these will be used later in further analysis.

Further information regarding avalanche occurrence is provided by figure 4.2 which shows a histogram of the number of avalanches occurring each day over the period 1974 to 2000. There is a definite split of data with many more avalanches recorded between 1974 and 1984 with the maximum number recorded on one day being 25. Avalanche observations are then lacking during 1985 and 1986, but then resume throughout the

1990's. This second period, however indicates a distinct reduction in avalanche occurrence with a maximum of 13 avalanches recorded on any one day during this time.

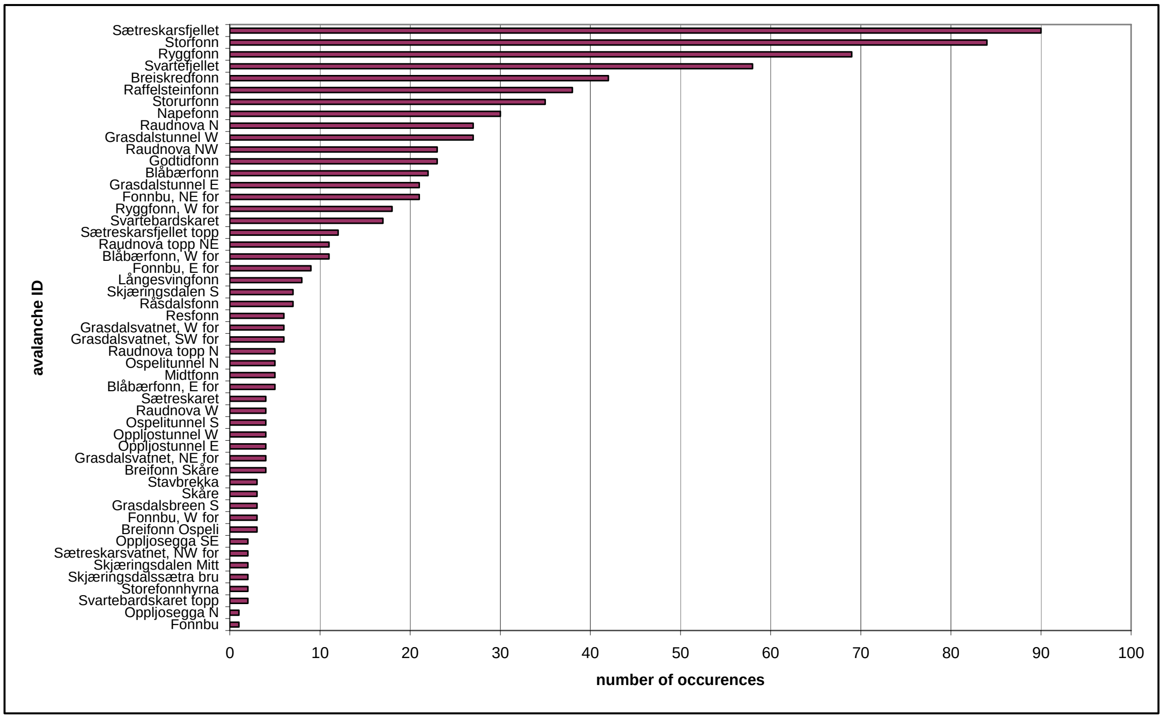


Figure 4.1: Graph showing the number of avalanches on each of the 51 paths between 1974 and 2000.

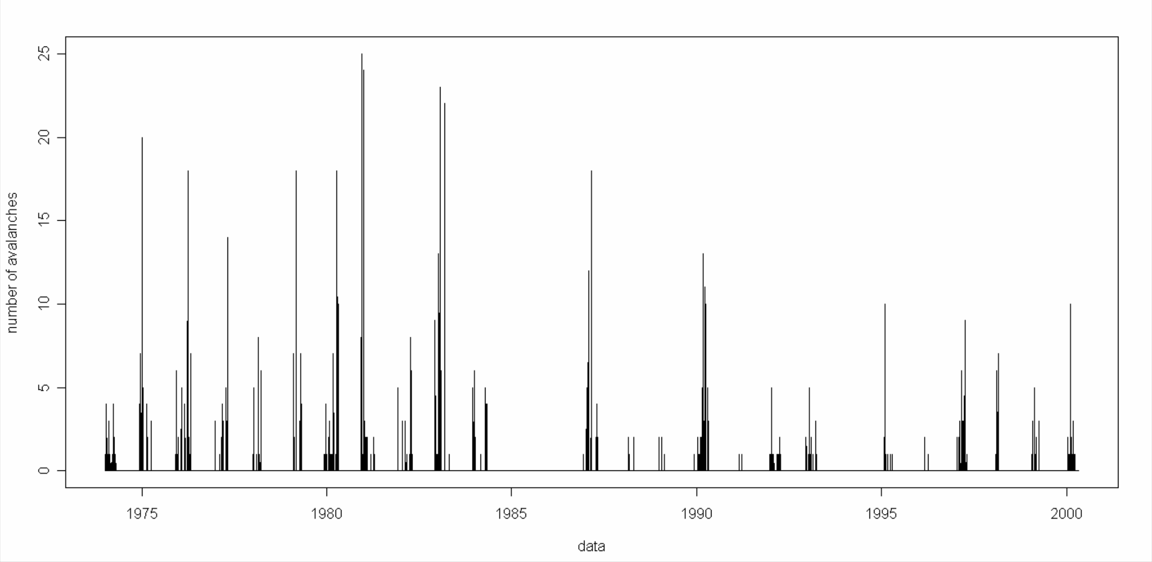


Figure 4.2: Histogram showing number of avalanches per day observed during the 25-year period

In addition to this initial sorting of avalanche data, statistical moments were calculated for certain meteorological parameters measured over the 25 year time period on avalanche days only. Table 4.2 presents a summary of these, with the mean and standard deviation given for the gridded data and also the Fonnbu data on avalanche days only. It is possible to see that there is great variation particularly for precipitation showing standard deviation approximately equal to the mean value for the majority of the precipitation (rr and RR) parameters. At present both wet and dry avalanche days have been included in the analysis, this may explain some of the large variation between results. In addition to looking at individual parameters in table 4.2, by comparing the mean values of precipitation in the gridded data with that of the Fonnbu data, large differences can be seen. The gridded data gives consistently much higher values of precipitation, approximately three times higher than that of the Fonnbu data.

gridded parameter code (all avalanche day data)	Mean	Standard deviation	Fonnbu parameter code (all avalanche day data)	mean	standard deviation
			RR0day	11	15
rr1day	40	40	RR1day	11	15
			RR2day	22	23
rr3day	84	71	RR3day	30	29
			RR4day	37	33
rr5day	116	91	RR5day	43	37
tam	-1.4	3.5	TAM1day	-3.9	4
wndspd1day	7.7	3.5	FFM1day	6.9	4.6
			FFX1day	10.7	6.9
			FXM1day	9.4	4.5
wndspd3day	7.8	2.7			
wndspd5day	7.5	2.4			
wndspdmax1day	9.7	3.8	FXX1day	15.9	9.8
wndspdmax3day	11.9	3.6			
wndspdmax5day	12.7	3.4			
			FXXmax1day	19.5	12.3
			FXXmax2day	21.6	13.9
			FXXmax3day	22.5	14.1
			FXXmax4day	23.2	14.1
			FXXmax5day	24.5	15

Table 4.2: Summary of statistical moments for all avalanche day data with precipitation measured in mm, temperature in ˚C and wind speed in m/s.

As wind speed may become an important factor with regards to the wind drift parameters at a later stage in this thesis, box-plots of the six gridded data wind related parameters are presented in figure 4.3. This gives a graphical representation of the order of magnitude of the wind speeds involved in this study. Here it is possible to see that an average wind speed of just over 5 m/s may be necessary in the run up to and on many avalanche days. In addition the mean maximum wind speed in the days preceding avalanche occurrence is between 10 and 12 m/s. There are however outliers, when wind speed has exceeded 20 m/s on some avalanche days, or hovered around zero on other days. These outliers are likely as, for example, on a day with almost zero wind, high levels of precipitation alone on this day may still give rise to avalanche conditions.

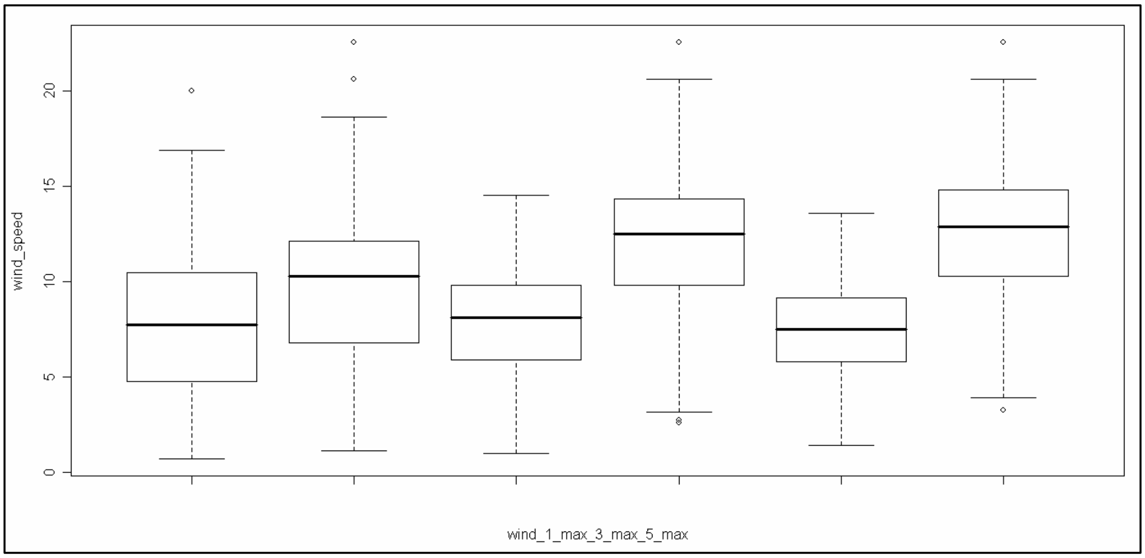


Figure 4.3: Box-plot of parameters wndspd1day, wndspdmax1day, wndspd3day, wndspdmax3day, wndspd5day, wndspdmax5day (from left to right). Wind speed is measured in m/s.

Further analysis showed good correlation between the one, three and five day measurements of wind speed, but the highest correlation of 0.9 was shown between wind speed over the three and five day period.

Another factor to be considered at this stage is wind direction. As previously mentioned in section 2.2.3 a relationship may often exist between prevailing wind direction and aspect of the avalanche starting zone. Figure 4.4 indicates the mean wind direction over the one, three and five day periods prior to avalanche events. These results are from analysis of the gridded data as no wind direction data is available in the Fonnbu data set at the time of writing. It is possible to see that approximately 65 % of the time the wind blows from a SW / W direction on avalanche days. Having found this fact out, the R program was used to see if any correlation existed between the global exposition code giving the aspect of the avalanche events compared to wind direction. This was however not the case, with a maximum correlation of 0.1 shown for the global exposition code in relation to the wnddir3day parameter.

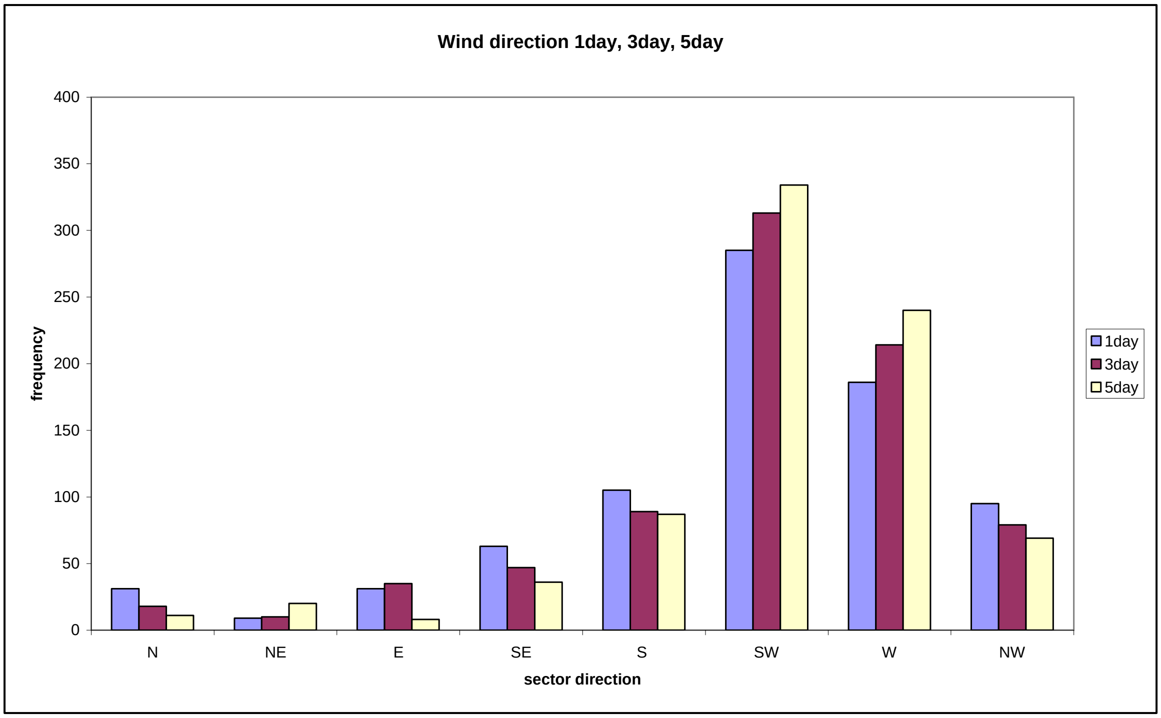


Figure 4.4: Histogram showing the dominant wind direction as SW to W, based on gridded wind data.

4.1.2 Dry avalanches only

The above section 4.1.1 has considered both the Fonnbu and gridded data irrelevant of the type of avalanches occurring. However, as an important part of this thesis is looking at the wind drift factor, the remainder of the exploratory data analysis shall be undertaken on just the dry avalanche data. This was done by removing the data rows which corresponded to days with wet avalanches. There were unfortunately a number of days when avalanches of unknown type were recorded, in this case, due to the large number of these it was decided to leave these within the data sets. In total there were 146 wet avalanches recorded in the gridded data which were subsequently discounted, and 68 days on which wet avalanches occurred in the Fonnbu data set that were also removed from further procedures.

Similar analysis was undertaken as in section 4.1.1 providing comparable results with only slight variation in the means of precipitation and wind speed data. Summary results are provided in table 4.3. The main difference to comment on was a 0.5˚C decrease in mean temperature in the Fonnbu data and 0.3˚C decrease in mean temperature in the gridded data when considering only days on which dry avalanches occur. This decrease was to be expected as temperature is a controlling factor in differentiating between wet and dry avalanches. A more surprising result when comparing the summary data in table 4.3 (dry avalanches) with that of table 4.2 (all avalanches) was that the range of values for each parameter is still very large. This is shown by the high values of standard deviation in relation to the mean values. Another point to note is that on comparing the gridded and Fonnbu mean parameter values of precipitation for the dry avalanche days only in table 4.3, although there is still a large difference, this has decreased slightly from the three fold difference noted in table 4.2.

gridded parameter code (dry avalanches only)	mean	Standard deviation	Fonnbu parameter code (dry avalanches only)	mean	standard deviation
			RR0day	11	15
rr1day	35	32	RR1day	11	15
			RR2day	23	23
rr3day	75	64	RR3day	31	30
			RR4day	39	34
rr5day	107	86	RR5day	45	38
Tam	-1.7	3.3	TAM1day	-4.4	4
wndspd1day	7.8	3.6	FFM1day	6.5	4.5
			FFX1day	10.5	7.3
			FXM1day	9.3	4.5
wndspd3day	7.8	2.7			
wndspd5day	7.4	2.4			
wndspdmax1day	9.8	3.9	FXX1day	16.1	10.2
wndspdmax3day	11.8	3.6			
wndspdmax5day	12.5	3.5			
			FXXmax1day	19.4	12.4
			FXXmax2day	21.3	13.7
			FXXmax3day	22.1	13.4
			FXXmax4day	22.6	13.2
			FXXmax5day	24.1	14.2

Table 4.3: Summary of statistical moments for only dry avalanche data with precipitation measured in mm, temperature in ˚C and wind speed in m/s.

4.1.3 Probability plot results

Following the above preliminary analysis, probability plots were then created using the dry avalanche data sets. This was undertaken for several precipitation parameters and combined wind drift parameters.

4.1.3.1 Using precipitation parameters

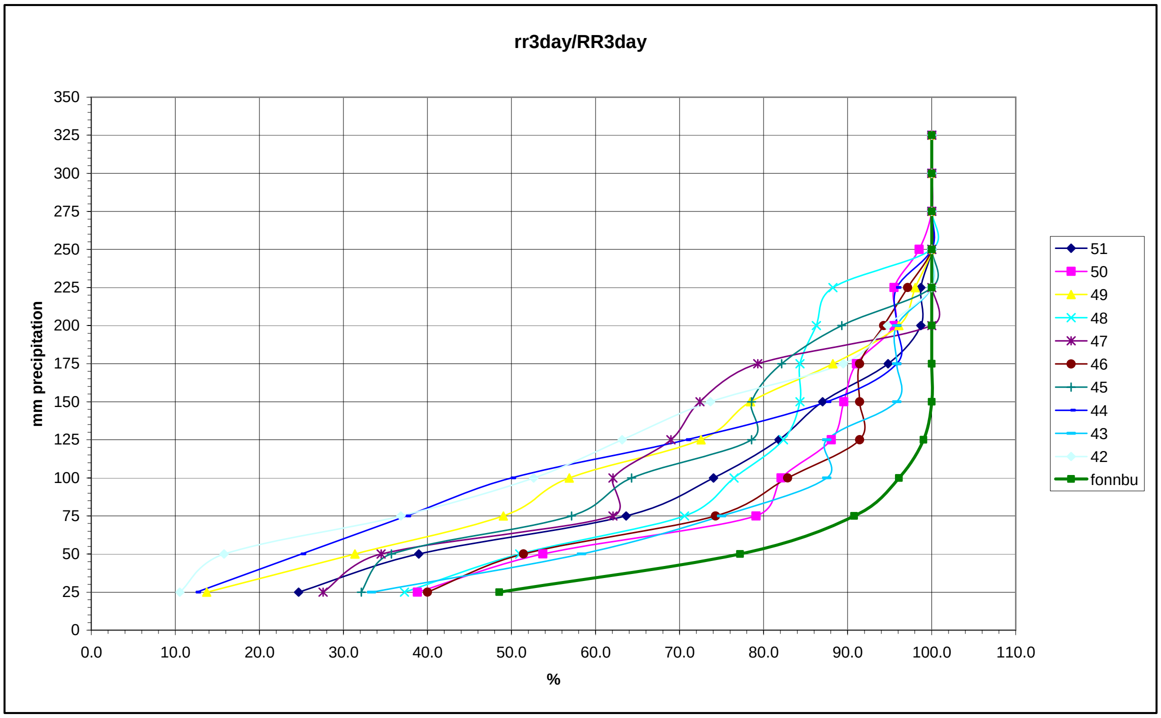


Figure 4.5(a): Cumulative probability plot of the three day sum of precipitation. Shown are the curves for the top ten most frequently occurring avalanches 51-42 from the gridded data set in addition to the thick green line of the Fonnbu data set.

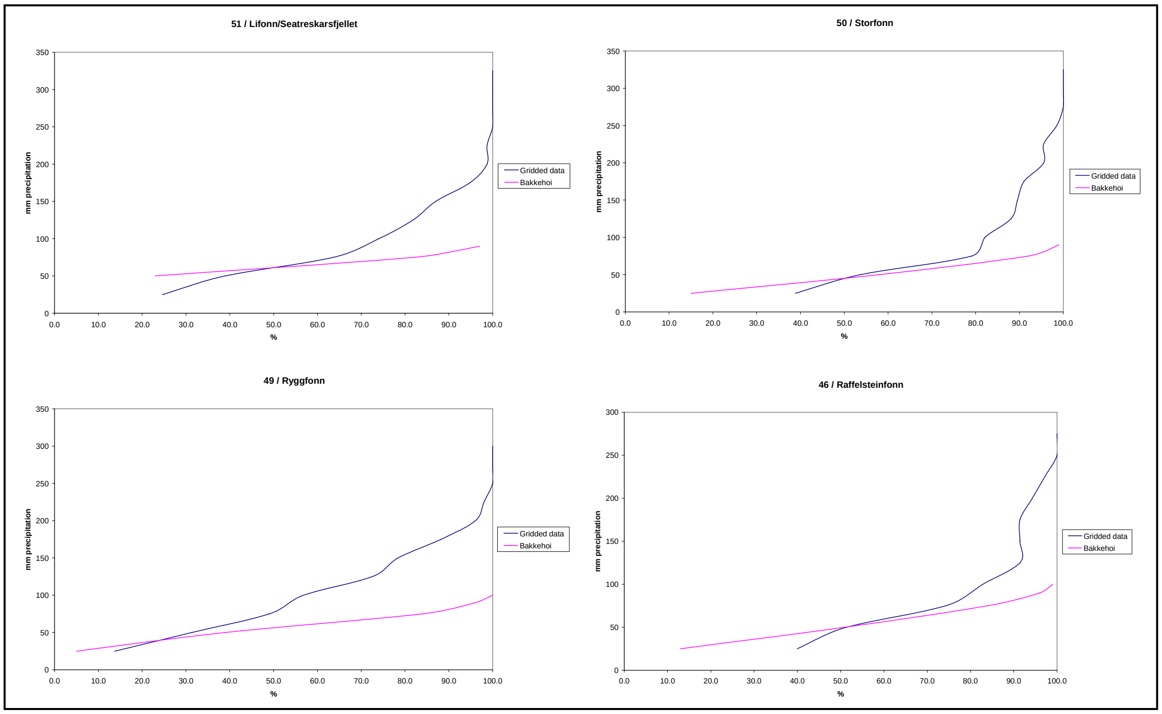


Figure 4.5(b): Individual probability plots for the three day sum of precipitation with gridded data shown by blue and Bakkehøi (1987) plots shown in pink. Clockwise from the top left for paths; 51 (Sætreskarsfjellet / Lifonn), 50 (Storfonn), 46 (Raffelsteinfonn) and 49 (Ryggfonn).

Figure 4.5(a) shows the cumulative probability plot of the three day sum of precipitation. This was chosen in order to make some comparison with a similar plot created by Bakkehøi (1987) which was previously mentioned in section 2.2.1. In figure 4.5(a), the top ten most frequently occurring avalanches were highlighted from the gridded data set. In addition, as the Fonnbu data set does not distinguish between avalanche paths, just the number of avalanches per day, this curve has also been displayed (in green) for comparison. It is possible to see that the curve depicting the Fonnbu data shows the lowest values for precipitation, in relation to probability, by a significant margin. This can be illustrated with the probability of avalanching at 80 %, the threshold of the three day sum of precipitation for the Fonnbu data set is 50 mm whereas the gridded data requires a threshold of between 75 mm and 175 mm depending upon the avalanche path.

Figure 4.5(b) shows probability plots for four of the avalanche paths taken from the larger plot of 4.5(a), with the addition of the corresponding probability plots for each path by Bakkehøi (1987). It should be mentioned here that the plot is a curved form as opposed to the linear relationship presented by Bakkehøi (1987), this is due in part to the scaling of the axis. Additionally the plots depicting results by Bakkehøi (1987) are only approximate values, having been interpreted from figure 2.1. These four plots in figure 4.5(b), in general, show a relatively close match between gridded data sets and the data set used by Bakkehøi (1987) for probabilities of around 50 %. However, for greater probabilities, much larger measures of precipitation are required for the gridded data.

From the probability plots of the three day sum of precipitation in figure 4.5(a), and additionally, plots of the one and five day sums of precipitation (see Appendix 2) it is possible to deduce threshold values. These are provided below in table 4.4 for the probability of avalanching at 50 %:

Gridded data		**Fonnbu data**	
parameter code	threshold values (mm)	parameter code	threshold values (mm)
rr1day	20-40	RR1day	5
rr3day	40-100	RR3day	25
rr5day	55-145	RR5day	40

Table 4.4: Showing threshold values at the 50 % probability level of avalanche occurrence from probability plots for the one, three and five day sums of precipitation.

4.1.3.2 Using combined wind drift parameters

Probability plots were created for the wind drift parameter over the one day, three day and five day period for the gridded data. This was done using the product of the precipitation and the corresponding mean or maximum wind speed to the fourth power as indicated in table 3.1(a). A total of six probability plots were therefore created, in all plots the axis was scaled for better visual interpretation. This often resulted in excluding the 100 % probability, as in the majority of cases this required a much higher wind drift factor.

Results varied between the ten probability curves, however, the overall trend was similar for each. Probability curves were also added from the Fonnbu data set for comparison. This was done for the wind drift parameter RRFFM1 which is comparable to the gridded data parameter 'rrwndspd1' and also RRFXX1 which can be considered comparable to the gridded data parameter 'rrwndspdmax1'. Figures 4.6(a) and 4.6(b) show these plots calculated using the mean wind speed values over the one day period for 4.6(a) and highest maximum wind speed values over the one day period in 4.6(b).

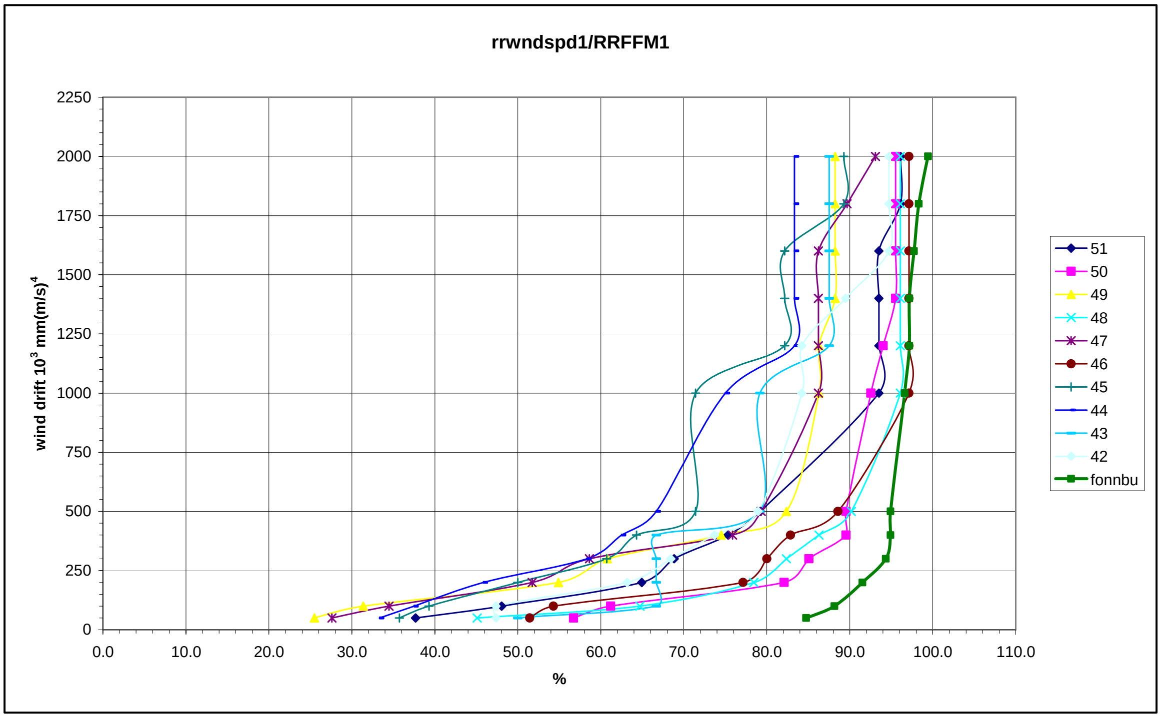


Figure 4.6(a): Cumulative probability plot of the wind drift factor using daily mean wind speed on the day of the avalanche. Shown are the curves for the top ten most frequently occurring avalanche paths 51-42 from the gridded data set in addition to the thick green line of the Fonnbu data set.

It is possible to see in figure 4.6(a) that the Fonnbu curve indicates the lowest values, however figure 4.6(b) shows this curve amongst the other ten probability plots. In terms of probability, at the 70 % avalanche probability level, figure 4.6(a) shows the wind drift threshold to vary between about 100 x 10^3 mm(m/s)^4 for paths 48 (Svartefjellet) and 50 (Storfonn) and up to 750 x 10^3 mm(m/s)^4 for path 44 (Napefonn), this is excluding the Fonnbu data results. This range doubles in figure 4.6(b) which relates to maximum wind speed values, whereby, at the same probability of 70 % the lower threshold has increased to 200 x 10^3 mm(m/s)^4, and the upper threshold is 1500 x 10^3 mm(m/s)^4.

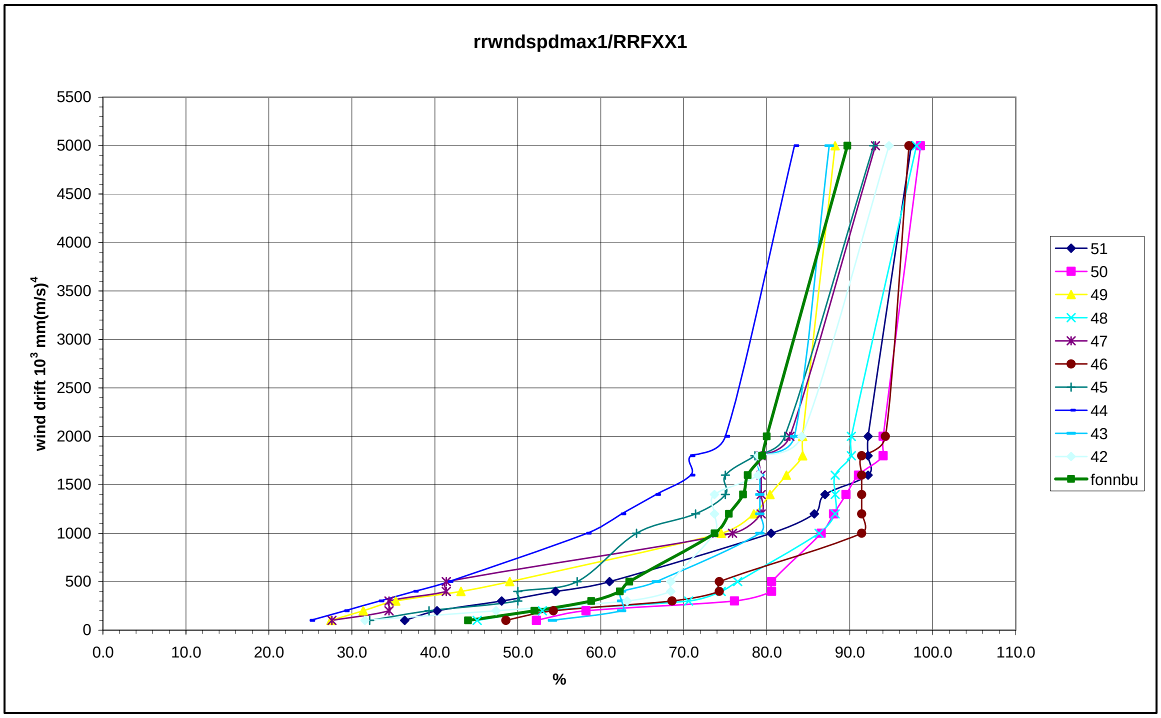


Figure 4.6(b): Cumulative probability plot of the wind drift factor using the highest maximum wind speeds on the day of the avalanche. Shown are the curves for the top ten most frequently occurring avalanche paths 51-42 from the gridded data set in addition to the thick green line of the Fonnbu data set.

4.2 Further statistical analysis

Only the Fonnbu data set was used from this point onwards in order to compare weather data recorded on days of avalanches with that on non avalanche days.

4.2.1 The Kruskal-Wallis results

The first procedure undertaken as outlined in section 3.4.3 was the Kruskal-Wallis test. The results of this are provided in table 4.5 which shows parameters in bold with significant difference in values, on avalanche versus non avalanche days, at the 5 % significance level ($p \leq 0.05$).

Parameter	p-value	Combined parameter	p-value
RR0day	**<0.001**	**RRFFM0**	**<0.001**
RR1day	**<0.001**	**RRFFM1**	**<0.001**
RR2day	**<0.001**	**RRFFX0**	**<0.001**
RR3day	**<0.001**	**RRFFX1**	**<0.001**
RR4day	**<0.001**	**RRFXM0**	**<0.001**
RR5day	**<0.001**	**RRFXM1**	**<0.001**
TAM0day	**<0.001**	**RRFXX0**	**<0.001**
TAM1day	0.069	**RRFXX1**	**<0.001**
TAN0day	**<0.001**	**RRFXXmax1**	**<0.001**
TAN1day	0.083	**RRFXXmax2**	**<0.001**
TAX0day	**<0.001**	**RRFXXmax3**	**<0.001**
TAX1day	0.11	**RRFXXmax4**	**0.0064**
SS0day	0.43	**RRFXXmax5**	**0.016**
SS1day	0.4		
SSdif1day	0.19		
SSdif2day	0.094		
SSdif3day	**0.008**		
SSdif4day	**0.0037**		
SSdif5day	0.081		
FFM0day	0.28		
FFM1day	0.44		
FFX0day	0.56		
FFX1day	0.3		
FXM0day	0.12		
FXM1day	0.17		
FXX0day	0.096		
FXX1day	**0.018**		
FXXmax1day	**0.0069**		
FXXmax2day	**0.016**		
FXXmax3day	0.21		
FXXmax4day	0.44		
FXXmax5day	0.32		

Table 4.5: P-values from the Kruskal-Wallis test for each parameter to indicate if a significant difference exists between data values on avalanche days compared to data on non avalanche days. Parameters in bold are significantly different ($p \leq 0.05$).

It is evident from the above table (4.5) that all the precipitation days (RR) show a significant difference between avalanche and non avalanche day values, this is also true for all the combined wind drift parameters. The other parameters show varying degrees of difference, for example the maximum highest maximum wind speed recorded over just the one or two day period shows a significant difference, however over a larger time scale of the three to five day sum this parameter shows little difference between avalanche days and non avalanche days. Additionally, temperature related parameters show a significant difference in temperature on the preceding day, however no significant difference is seen in temperature on the actual avalanche or non avalanche day.

4.2.2 Classification tree results

In accordance with section 3.4.4.2, fifty classification trees were created using the R language and environment (R Development Core Team, 2005). The same set of data was used each time for the dry avalanche days, but a random set of non avalanche days were picked for each of the runs. Graphical representation of each tree is provided in Appendix 4, however a summary of results are given in table 4.6 showing the run number alongside the parameters used in the first three splits of the tree, ranked in order of importance from 1 to 3. In addition threshold values are shown alongside each corresponding splitting parameter. Combined wind drift parameters are shown in bold and it is possible to calculate that of the 50 runs only 16 of these i.e. 32 % provide classification trees with any of the 13 combined wind drift parameters ranked within the top three splits.

The most significant result to note from the 50 runs is that the first split is dominated by precipitation parameters, particularly RR2day and RR4day which account for almost half of the cases in split 1. The other parameters present in split 1 consist of the remaining precipitation parameters and several of the combined wind drift parameters. Splits 2 and 3 show more variation across the different parameters, with a total of 16 different parameters appearing in split 2, and 19 different parameters in split 3. These second and third place splits are however, much more dominated by temperature related parameters, particularly that of TAX0day which is the most recurring parameter across all of the trees.

Run	Split 1	Thres-hold	Split 2	Thres-hold	Split 3	Thres-hold	% Misclass.
29	RR2day	>7.25	**RRFXX1**	>1.92e5	TAX0day	>4.7	19.6
10	RR4day	>30.3	SSdif2day	<5.5	TAX0day	>4.7	19.6
19	RR5day	>37.85	TAN1day	>-10.95	TAX0day	>4.7	19.6
31	RR2day	>20.8	TAX0day	>6.15	FXX0day	>17.45	19.9
13	RR5day	>32.45	TAX0day	>4.6	FXX0day	>12.5	20.6
41	RR2day	>20.85	TAM0day	>-0.44	RR3day	>9	21.3
24	RR2day	>24.35	TAM1day	>-5.01	**RRFXM1**	>1.32e4	21.7
42	RR4day	>32.9	TAX1day	<1.55	TAM1day	>-5.18	21.7
47	RR3day	>26.55	TAX0day	>4.95	FXX0day	>12.4	22.0
3	RR3day	>28.95	TAM0day	>-0.4	**RRFXXmax1**	>1.45e6	22.4
1	RR2day	>19.25	TAX0day	>4.7	RR5day	>9.35	22.7
37	RR5day	>37.85	FXX0day	>21.2	FFM1day	<3.75	22.7
14	**RRFXXmax1**	>4.35e5	TAM0day	>-0.43	RR5day	>42.85	23.1
21	RR3day	>35	FXX0day	>12.4	FFM0day	<6.88	23.4
38	RR4day	>32.9	TAM0day	>-3.24	**RRFXXmax1**	>5.76e5	23.4
2	RR2day	>8.9	TAM1day	<-0.14	TAX0day	>4.75	24.1
36	RR2day	>25.45	TAM0day	>-3.24	FXXmax1day	>21.5	24.1
8	RR3day	>34.45	FXX0day	>24.1	TAM0day	>-0.46	24.1
26	RR4day	>32.9	TAX0day	>4.75	FXXmax1day	>15.1	24.1
7	RR5day	>23.85	TAX0day	>4.75	FXX0day	>16.35	24.1
12	RR2day	>20.85	TAN0day	>-0.15	RR3day	>7.65	24.5
44	RR2day	>19.5	TAN0day	>-10.1	FXXmax1day	>21.2	24.8
49	RR4day	>32.4	FXX0day	>24.1	TAN1day	>-8.95	25.2
4	**RRFXX1**	>2.17e5	TAX0day	>4.75	RR4day	>32.6	25.2
35	RR0day	>4.05	TAM0day	>-0.4	FXX0day	>24.1	25.5
17	RR4day	>32.6	**RRFXX1**	>1.97e5	TAX0day	>4.75	25.5
40	RR4day	>34.2	FXX0day	>16.35	TAX0day	>4.75	25.5
27	RR5day	>37.35	TAN1day	>-9.2	SS1day	<226.5	25.5
48	RR2day	>19.4	TAM0day	>-3.59	FXM1day	>2.83	25.9
22	RR3day	>32.3	**RRFXX0**	>1.79e5	TAX0day	>4.75	25.9
46	RR3day	>28.95	FXXmax2day	>21.2	FFM1day	<19.83	25.9
39	RR4day	>32.9	FXXmax1day	>21.2	TAX0day	>4.45	25.9
28	**RRFXX0**	>1.66e5	TAX0day	>4.7	RR4day	>32.4	25.9
45	RR3day	>25.85	TAX0day	>4.7	FXXmax4day	>23.95	26.6
25	RR5day	>30.3	FXX0day	>24.1	TAX0day	>4.7	26.6
33	RR5day	>30.3	FXX0day	>16.35	FXX0day	<18.5	26.6
23	**RRFXX0**	>1.63e5	RR5day	>31.35	TAX0day	>4.55	26.6
16	RR2day	>15.25	FXXmax1day	>21.2	TAM1day	>2.36	26.9
9	RR4day	>29.7	TAX0day	>-3.45	FXX0day	>21.2	26.9
32	RR2day	>9.15	TAN1day	>-10.45	TAM0day	>-0.34	27.3
6	RR4day	>35.15	FXX0day	>16.4	TAX0day	>5.55	27.3
15	RR4day	>14.8	TAM1day	>-2.56	RR5day	<3.45	27.3
50	RR5day	>42.9	TAM0day	>1.25	**RRFXX1**	>2.03e5	27.3
43	**RRFXXmax1**	>7.89e5	RR3day	>35	TAM0day	>2.45	27.3
30	RR3day	>9.05	TAM1day	<-0.14	TAX0day	>4.7	27.6
20	**RRFXXmax2**	>1.06e6	TAM1day	>1.06	RR5day	>30.1	27.6
5	RR5day	>29.65	TAM0day	>0.45	**RRFXXmax4**	>5.47e6	28.3

11	RR4day	>35	FXX0day	>24.1	TAX0day	>6.1	28.7
18	**RRFXXmax1**	>4.16e5	SSdif4day	<-51.5	TAX1day	>5.2	28.7
34	**RRFXX1**	>1.90e5	**RRFXXmax5**	>1.51e7	TAX0day	>4.95	29.0

Table 4.6: Summary of results for 50 classification trees showing parameters selected for the 1st 2nd and 3rd place splits along with their threshold values and the overall percentage misclassification for each tree. Wind drift parameters are indicated in bold.

It is possible to summarize table 4.6 to show the most frequently recurring of the 45 parameters within the top three splits of the 50 trees. These parameters, recurring five or more times, are shown in table 4.7. Excluded from this are 18 parameters that occur infrequently and 17 parameters that do not occur at all within the top three splits. From table 4.7 it is possible to deduce that overall, precipitation parameters are the most important factors, closely followed by temperature related factors, beyond this, wind and combined wind drift parameters feature.

parameter	**# in 1st place**	**# in 2nd place**	**# in 3rd place**	**total**	**average threshold value**
TAX0day	0	10	14	24	>4.52 ˚C
FXX0day	0	9	7	16	>18.85 m/s
RR4day	12	0	2	14	>31.5 mm
RR5day	9	1	4	14	>30 mm
RR2day	12	0	0	12	>17.6 mm
TAM0day	0	9	3	12	>-0.7 ˚C
RR3day	8	1	2	11	>24.8 mm
TAM1day	0	5	2	7	>-1.41 ˚C
FXXmax1day	0	2	3	5	>20 m/s
RRFXX1	2	2	1	5	>2.0e5 $mm(m/s)^4$
RRFXXmax1	3	0	2	5	>7.3e5 $mm(m/s)^4$

Table 4.7: Showing a summary of the most frequently occurring parameters within the top three splits of the 50 classification trees. The greater than (>) sign signifies that dry avalanches occur when the parameters exceed the stated threshold.

In addition to the classification trees showing the parameter selected at each split, they also provide threshold values. These are shown for each of the splitting parameters in table 4.6, but have been summarized in table 4.7 whereby an average for each frequently used parameter has been calculated. Average threshold values for precipitation vary between approximately 17 to 30 mm over the varying time periods. At first glance TAX0day shows a surprising result with a positive average threshold value of 4.52˚C i.e. dry avalanches occur when TAX0day exceeds this temperature. The implications of this will be discussed in detail later in section 5.4. It can also be seen from table 4.7 that maximum wind speed parameters, with values fluctuating between 18 and 20 m/s, are frequently observed as important factors for tree splitting to distinguish between dry avalanche and non avalanche days. Finally, in terms of the combined wind drift factors the ones to note are the two parameters RRFXX1 and RRFXXmax1 which use maximum wind speeds on the day of the avalanche. Here, threshold values vary between 200,000 and 730,000 mm(m/s)^4.

Besides selecting the parameters and threshold values for each split, each classification tree had a misclassification rate. In table 4.6 the percentage misclassified refers to the total number of misclassified events at the end nodes, that is to say the number of non avalanche days classified as dry avalanche days and vice versa, this is calculated as a percentage of the total number of events, i.e. 286 dry avalanche and non avalanche days. The classification trees have been created using pruning conditions giving generally only enough complexity in order to analyse the top splits of the tree, as this is of greater relevance in this thesis. Due to this the trees generally provided greater misclassification rates than would be suitable for forecasting purposes (Föhn, 1998 in Kronholm *et al*., 2006a), with misclassification varying between 20 and 30 %.

Classification trees 4 and 29 have been included in figures 4.7(a) and (b) respectively as examples of possible formats. They are both pruned using the same complexity parameter of 0.03, but as tree 29 in figure 4.7 (b) uses more splitting parameters this has the lower misclassification rate of 19.3 % compared to tree 4 in figure 4.7 (a) with a misclassification rate of 25.2 %. The main point to note with these trees, however, is that although they look very different, the parameters used within the first three splits are very similar, although in a different order of importance:

Split 1	RRFXX1 (Run 4)	RR2day (Run 29)
Split 2	TAX0day (Run 4)	RRFXX1 (Run 29)
Split 3	RR4day (Run 4)	TAX0day (Run 29)

And with the exception of the precipitation parameters (RR2day and RR4day) they have very similar threshold values.

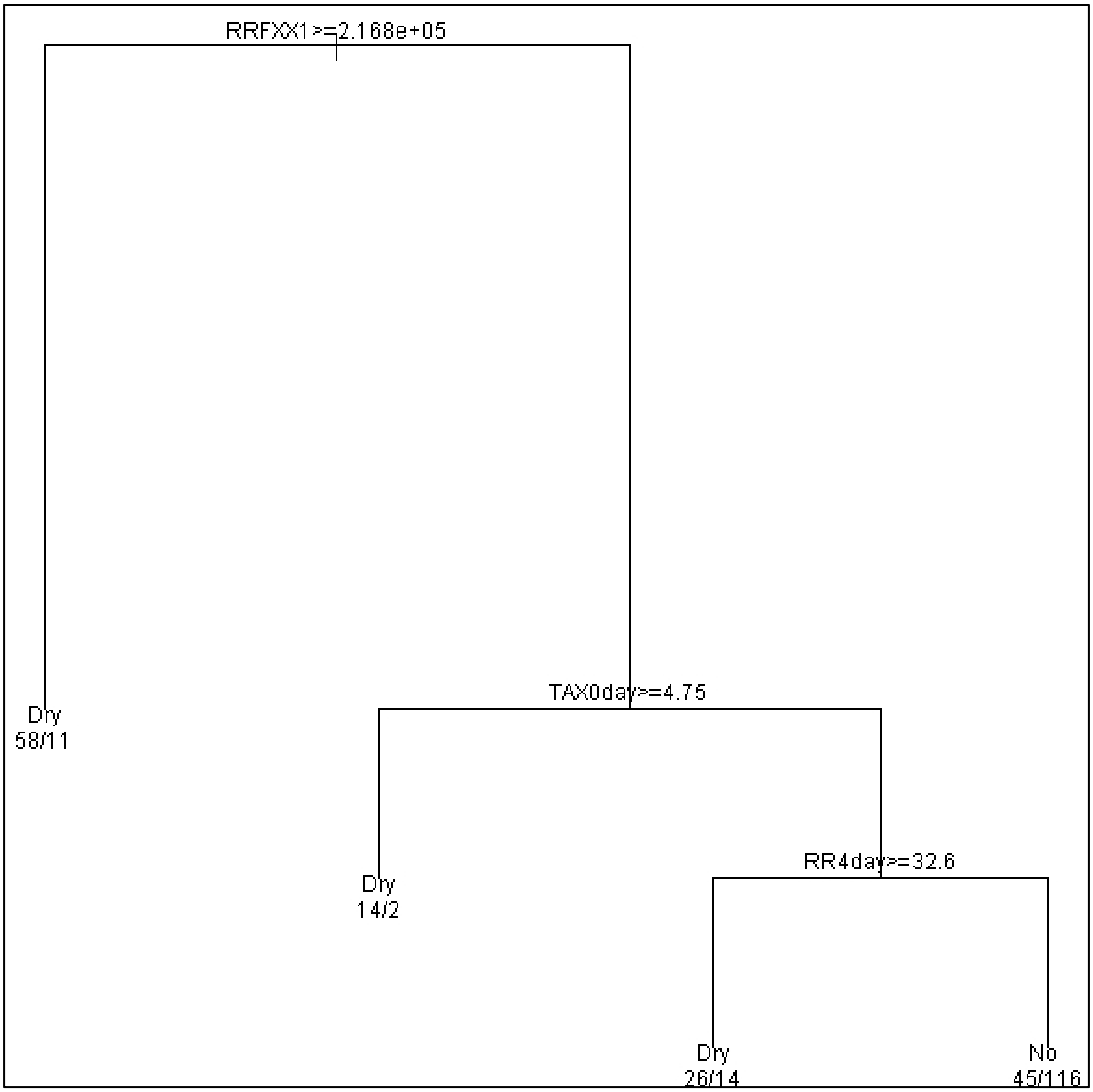


Figure 4.7 (a): Classification tree 4 with an overall misclassification rate of 25.2 %.

Both figures (a) and (b) are grown using the random number of non avalanche days = the number of dry avalanche days = 143. At each end node the term ‘No’ refers to days classified as non avalanche days and ‘Dry’ refers to the days classified as dry avalanche days. Below this, the two numbers depict the correct number of classified events at each end node and the number of misclassified events at each end node, by way of Ndry / Nno. The vertical length of each branch is proportional to the ability of each node to split correctly. This explanation is in-line with that by Kronholm *et al*. (2006a).

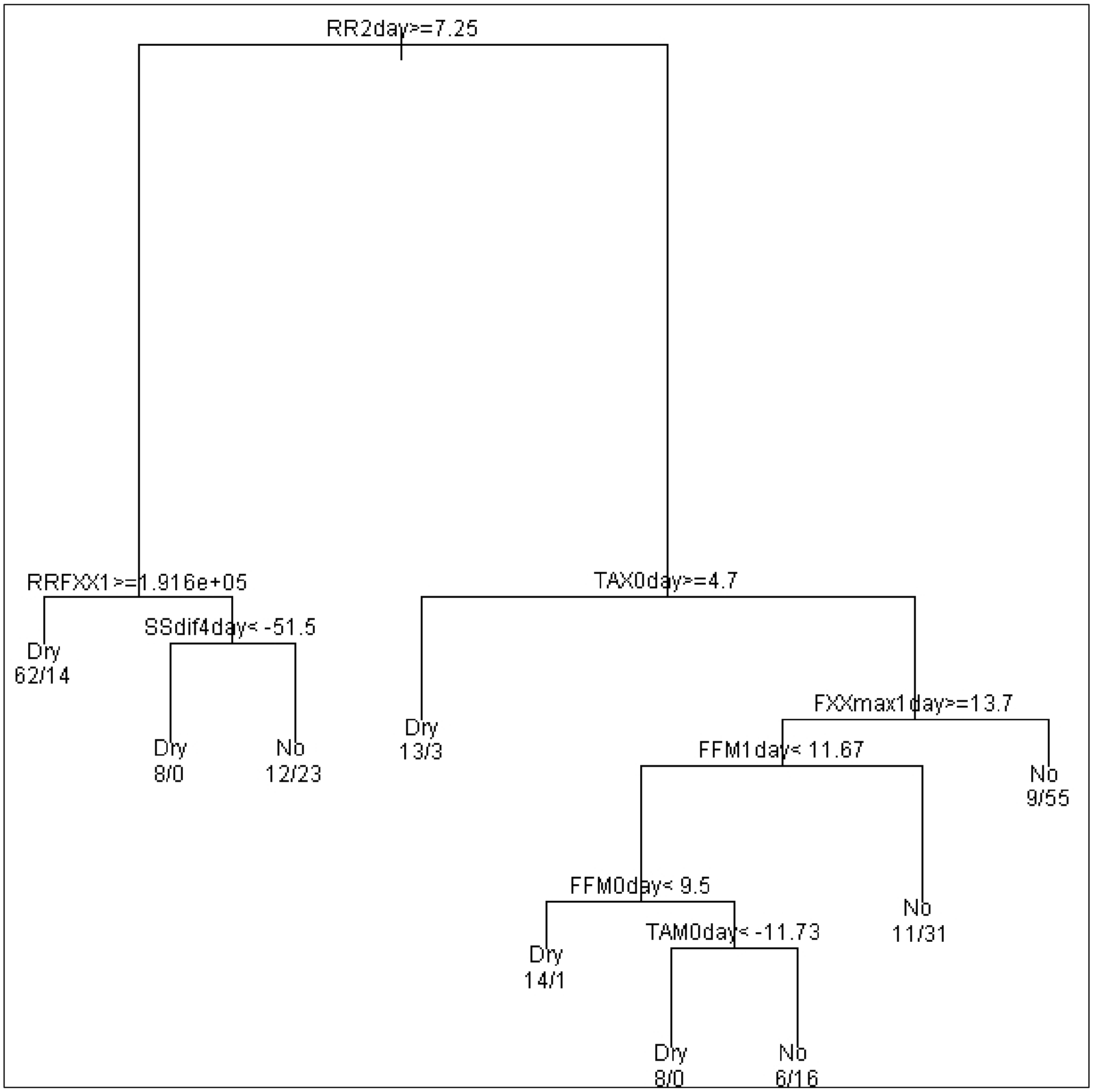


Figure 4.7 (b): Classification tree 29 with an overall misclassification rate of 19.6 %.

It is also apparent in figure 4.7(b) that along with the top three splits previously discussed, a varied number of the other parameters may also occur lower down the tree. As the combined wind drift parameters are a fundamental part of the data analysis in this thesis, it should be mentioned that although they may not seem as significant as other parameters within the top three splits, overall, of the 50 trees created, 21 trees used at least one wind drift parameter at one of the splits in the pruned result. There is however great variation in placement of these combined parameters from the top splitting node 1 down to the lowest nodes.

This chapter has provided a summary of results from initial exploration of all the data sets, to statistical procedures with finally the creation of a number of classification trees using only selected data from the Fonnbu data set. These results will be commented on further in chapter 5 and compared with published results in the discussion sections of this thesis.

Chapter 5: Discussion

5.1 Appraisal of the data sets

To summarise, the results in this thesis have been based on four sets of data from a variety of sources with differing degrees of accuracy and frequency of measure. Three of the data sets have been combined to form an interpolated gridded set and one set are actual observed results from the Fonnbu site in Grasdalen. These two sets will be discussed in section 5.1.1 below.

5.1.1 Gridded data versus observed Fonnbu data

Initial analysis was undertaken on both the gridded data and observed Fonnbu data, with the aim to see their similarities but also what differences exist. Immediately, it is possible to see there is large variation in precipitation values between the gridded data and Fonnbu data as presented in the summary tables 4.2 and 4.3. This may be explained by distortion of the gridded data being due to the use of an irregular network of weather stations located with decreasing density at higher elevations, in combination with an interpolation logarithm (Norwegian Meteorological Institute, 2004). In this thesis it is apparent that this distortion has led to overestimation, with much higher values for the one day, three day and five day sums of precipitation in the gridded data compared to the Fonnbu data set.

Another reason for the difference in precipitation is the obvious fact that, although the gridded data is extrapolated, it is still based on observed measurements. The fact that the study site is in a mountainous environment tends to mean that there are significant local and micro-scale wind effects (Barry, 1981), and certain areas experience much more localised precipitation than other sites. It is also important to mention that, under the influence of wind, precipitation gauges can severely underestimate precipitation as snowfall (e.g. at the Fonnbu site), compared to rainfall (e.g. at lower elevation weather stations). This indicates how variation can exist in precipitation measurements at different sites and altitudes dependent upon the form of precipitation and wind conditions. Additionally, uncertainties and errors between the data sets can arise due to differing types of precipitation gauges which may be in use and the quality and quantity of measurements taken at each of the sites. All of these factors could mean that the initial measurements from weather stations used for the gridded data may indicate some overestimates before the interpolation logarithm is even applied.

It must however be mentioned that although there is a large difference in mean measurements of precipitation between the gridded and Fonnbu data, the range of this data is also very large. Standard deviations are shown to be approximately equal to the mean values in each of the data sets, this indicates the vast range in precipitation measurements, all of which can be influential in avalanche formation.

In addition to precipitation differences between the gridded and Fonnbu data there are also differences in wind speed, these are seen to fluctuate, with the gridded data giving higher values for the average wind speeds on the day of the avalanche. However the Fonnbu data shows higher values for maximum wind speeds on the day of the avalanche. One reason for this difference between data sets could be that they are not directly comparable as has been implied by tables 4.2 and 4.3.

A final point to note is the difference in temperatures, the Fonnbu data sets shows mean daily temperatures several degrees lower than those of the gridded data. This is in agreement with Kronholm *et al*. (2006a) whereby the gridded data does not fully reflect the large temperature gradients often apparent due to the large elevation differences within the area. Also significant to mention is the comparison of mean temperatures across all avalanche days (table 4.2) with just dry avalanche day results (table 4.3). The slight decrease in temperature on dry avalanche days was unsurprising and shows that this factor has some bearing on avalanche type and formation.

The above discussion helps explain the choice to use the Fonnbu observed data for the further statistical analysis stages in this thesis, as overall they can be considered more accurate for just the localised Grasdalen area. In addition, statistical procedures such as classification trees with the use of the combined wind drift factor have not previously been carried out on the Fonnbu data set.

5.1.2 Avalanche data

With regards to avalanche day recording throughout the data sets, there is some variation apparent in recording of avalanche days over the 25 year period. This is evident from figure 4.2 in which the first ten years (1974 to 1984) show a consistently greater number of avalanches being recorded than the more recent decade in the 1990's. Reasons for this decrease in observations may be that there were actually just fewer avalanches occurring or alternatively, avalanches were not as accurately recorded in recent years compared to when records were first started. Also with greater mitigation and hence controlled release of avalanches there are less naturally occurring avalanches.

It has been mentioned that often subjective observations leading to inaccurate records of avalanche type may be partly to blame, as what are actually wet avalanches may often be wrongly recorded as dry avalanches and vice versa, and additionally avalanches of unknown type have been recorded. Nothing can be done to reduce this error as disqualifying all avalanches of unknown type would reduce the data set too much.

Finally, to increase accuracy of the data set, the avalanches with a large uncertainty in time scale of greater than +/- 12 hours were disqualified from further analysis. This was, however, one of the few ways to manipulate and limit the errors and uncertainties in the data sets.

5.2 Assessment of important weather parameters

A number of different weather parameters were provided within the data sets the most important ones in terms of this thesis are considered below.

5.2.1 Precipitation

Precipitation has long since been cited as one of the single main reasons for avalanche formation (Zingg, 1965; Akkouratov, 1965; de Quervain, 1965; Bakkehøi, 1987). This fact has been highlighted in this thesis with the in-depth analysis of various parameters of precipitation. These have included the one day, three day and five day sums from the gridded data, and the preceding day, one, two, three, four and five day sums from the Fonnbu data. Probability plots were created for the one, three and five day parameters, with threshold values provided in table 4.4. From the probability plots (figure 4.5(a) and Appendix 2) it is apparent that the five day sum shows a slightly better match between the Fonnbu and gridded data, although all plots show the Fonnbu data with the lowest measure of precipitation in relation to the highest probability of avalanching.

On excluding the Fonnbu results on these probability plots it becomes apparent that path 42 (Grasdalstunnel W) is one of the least likely to avalanche, this is expected as, of the ten paths chosen, this path is the least frequent to avalanche. Slightly more surprising is the result that path 51 (Sætreskarsfjellet), the most frequent path to avalanche is not considered the most probable or easiest to trigger. This is indicated in figure 4.5(a) which shows that it requires an approximate 60 mm sum of precipitation over the three day period to have a 50 % probability of failure. With the same precipitation, path 43 (Raudnova N) has a greater probability of 65 % to avalanche. One of the main reasons for this has been outlined by Bakkehøi (1987) who states that as path 51 (Sætreskarsfjellet) has a lower slope inclination than some of the other paths, then probability of failure will be lower for this path with small precipitation values. Another possibility may be that as path 51 (Sætreskarsfjellet) is the most frequent to avalanche it has most controlled releases during shorter periods of lower precipitation before snow levels can build up, but natural release avalanches then often occur during periods of heavier or lengthier precipitation when it is less possible to control the slope artificially.

The probability plots can also be compared with that of figure 2.1 created by Bakkehøi (1987) for the same location. This is shown in figure 4.5(b) for the paths; Raffelsteinfonn, Ryggfonn, Storfonn and Lifonn (also known as Sætreskarsfjellet) chosen for inclusion by Bakkehøi (1987), which relate to paths 46, 49, 50 and 51 respectively from the gridded data set. It is possible to see that at the 50 % probability level, for three of the four plots, precipitation levels are approximately equal between the gridded data and plots by

Bakkehøi (1987). At this level of probability, precipitation levels of approximately 50 mm are required on these three plots. For larger values of probability (greater than about 80 %) there is much greater difference between the gridded data and data used by Bakkehøi (1987). One reason for this could relate to the initial measurements of precipitation for the gridded data compared to the Fonnbu data. It has been seen in table 4.2 that on avalanche days there is a three fold increase between precipitation recorded in the gridded data set compared with that measured at the Fonnbu weather station. In relating this to the probability plots, a similar multiplying factor is seen at the 100 % probability level with plots by Bakkehøi (1987) showing levels of precipitation required at approximately 100 mm, whereas the gridded data shows requirements of 200 to 300 mm of precipitation.

5.2.2 Wind speed

As is evident from figure 4.3 and table 4.3 average wind speeds on avalanche days in both the gridded data and Fonnbu data fluctuate around 7 m/s, with maximum wind speeds over 20 m/s for the Fonnbu data on the days preceding avalanching. This is in line with suggestions by McClung and Schaerer (1993) who state that a minimum threshold of 4 m/s is necessary for snow transport and maximum snow transport is attained at wind speeds of about 20 m/s. Although from initial analysis high enough wind speeds are attained on avalanche days to produce significant snow transport, it would then seem to follow that on non avalanche days wind speeds may be significantly lower. Using the Kruskal-Wallis test this however does not hold true. With the exception of the three parameters of maximum wind speed measurements (FXX1day, FXXmax1day and FXXmax2day) which show significant differences between avalanche and non avalanche days, all the other wind speed parameters show no significant difference. This is in contrast to results found by Kronholm *et al*. (2006a; 2006b). Although they tested the gridded data set rather than the Fonnbu data set as used here, all the wind parameters were seen to show a significant difference between avalanche and non avalanche days. A further point to note here, is that in this thesis only one set of random non avalanche days were selected for the K-W test to compare with the avalanche days, had a different set of non avalanche days been selected this may have altered the results particularly considering that only 294 non avalanche days were chosen from a possible 3588 available non avalanche days.

5.2.3 Wind direction

Although only accounted for in the preliminary analysis of this thesis, wind direction is often considered to play an important role in avalanche formation (McClung and Schaerer, 1993). From figure 4.4 it can be seen that wind direction on avalanche days in Grasdalen is predominantly from the south west. This on first glance seems a simple deduction and in accordance with McClung and Schaerer (1993) and Pomeroy and Gray (1995) it therefore suggests that the mountainsides with a NE / E aspect are more prone to frequent avalanche occurrence. An example to highlight this refers to avalanche 51 (Sætreskaresfjellet), as previously mentioned it is the path most frequent to avalanche and additionally has an ENE aspect as would be expected. However, several problems exist with this interpretation, the first being that the gridded data may not reflect the true wind direction as it can be greatly modified on a local scale. The very heterogeneous topography of the area causes wind to be funnelled through valleys and gullies and re directed around rock outcrops (Barry, 1981). In addition to this factor, a number of avalanche paths are located in cirques, their recorded aspect may therefore vary by up to 180˚. This was evident as little correlation was seen between the majority of the recorded avalanche exposition codes in relation to wind direction.

No wind direction data was provided in the Fonnbu data set so no further comparisons are possible with observed data from the area. It is however likely that results from the Fonnbu site, being in a valley, would indicate that wind is funnelled into one of two broad directions.

5.2.4 Temperature

Temperature must be mentioned as this is seen to play a predominant role within the classification trees. The parameter outlining maximum temperature on the preceding day of the avalanche was the most important factor in both second and third place splits. The significance of temperature in this thesis is in agreement with Butler (1986) who states that of 223 avalanches in Glacier National Park, Montana, 80 % were related to changes in temperature. In this thesis, however, the importance of temperature was slightly unexpected particularly as dry avalanches only were considered for tree creation. Temperature is generally accepted to be of greater significance for wet avalanche occurrence. More will be mentioned regarding temperature at a later stage in this discussion.

5.3 Evaluation of combined wind drift data

Wind drift data was derived to create thirteen new parameters in the Fonnbu data set and six new parameters in the gridded data set.

5.3.1 Ease of application

Several of these combined wind drift parameters from both data sets were then used in the creation of probability plots. Additionally the Fonnbu combined wind drift parameters were applied in both the Kruskal-Wallis test and in the creation of classification trees. Although the creation of these parameters was simple enough, the data values are somewhat meaningless numbers, with threshold values that become more complex to interpret than the single parameter elements. As an example, from figure 4.6(b) it is possible to see at the 70 % probability, the threshold levels of the wind drift factor vary between 250,000 mm(m/s)^4 to 1,600,000 mm(m/s)^4. These values can be created from any number of combinations of precipitation and wind speed, it is therefore difficult to relate these back to the single parameter values they are derived from. For example, numerically:

$$\text{Precipitation x (wind speed)}^4 = 250{,}000 \text{ mm(m/s)}^4 \qquad \text{[Eq. 5.1]}$$

Where: precipitation = 1 mm and wind speed = 22.4 m/s

Or: precipitation = 104 mm and wind speed = 7 m/s

Although the result gives the same threshold value, would these two instances give rise to a realistic avalanche risk of 70 %? Perhaps it is necessary to look at threshold limits on the single parameters before calculation of the combined parameters, as SLF (2006) have previously stated wind transported snow reaches a maximum at wind speeds of between

15 – 20 m/s. With this in mind, under conditions of low precipitation but very high wind speeds probability of avalanching should decrease, this is not permitted in using the expression derived above [Eq. 5.1].

5.3.2 Comparison with published work

5.3.2.1 In terms of the Kruskal-Wallis procedure

Firstly it is appropriate to mention the significance of the wind drift factor by way of the Kruskal-Wallis test. This test has previously been carried out by Kronholm *et al*. (2006b) using approximately the same six wind drift parameters as proposed in this thesis created from the gridded data. All these parameters showed a large significant difference ($p \leq 0.001$). This is comparable with results of the Kruskal-Wallis test in this thesis which uses combined parameters from the Fonnbu data set. This shows all thirteen combined wind drift parameters having significantly different results between avalanche days and non avalanche days at the 5 % significance level, with the majority also at ($p \leq 0.001$).

The above result, although valuable in terms of the combined wind drift data, it should be mentioned that the Kruskal-Wallis test may not be the best procedure for this data set. This is due to the fact that some authors quote the Kruskal-Wallis test to be used for three or more samples (Graphpad, 1999; Wikipedia, 2007c; Wheeler *et al*., 2004) with the better test for two samples, as in this case of avalanche days versus non avalanche days, being the Mann-Whitney U test or Wilcoxon test (Wheeler *et al*., 2004). These tests were not considered for use for two reasons; the first being that the R program was to be used and documentation for this program states that the Kruskal-Wallis test is applicable for two or more samples (R Development Core Team, 2005). Also this test, being the same as used by Kronholm *et al*., (2006a) for the gridded data set, provides a fairer comparison between results.

5.3.2.2 In terms of the classification tree procedure

Both Davis *et al*. (1999) and Hendrikx *et al*. (2005) have used combined wind drift parameters to analyse avalanche occurrence. Hendrikx *et al*. (2005) created six combined parameters, these were over the 12, 24 and 72 hour period, with one set accounting for positive temperature readings by assigning these a value of zero wind drift, and the other set irrelevant to temperature. At the stage of deriving wind drift parameters in this thesis, wet avalanches were excluded, making the combined results to some extent comparable to the temperature dependent parameters of Hendrikx *et al*. (2005).

Davis *et al*. (1999) created a number of combined wind drift parameters, all of which take avalanche type into consideration as wet avalanche days are excluded. Three of their combined parameters are comparable with those created in this thesis, these include the use of total precipitation and wind speed over the 24, 48 and 72 hour period. Davis *et al*. (1999) have also created eight combined parameters categorising wind speeds into sectors of direction. This detail is not included in this thesis, primarily due to there being no data regarding wind direction in the Fonnbu data set, and additionally because this thesis is an initial investigation into the role of wind drift in Stryn. Classification trees created by Davis *et al*. (1999) used various combinations of data, some used only derived wind drift parameters, and these were compared with trees created using all the weather parameters. This was not the case for this thesis, as trees were created by submitting all weather parameters. It was thought this would be more suitable at this stage to provide a more robust procedure and highlight the most important splitting factors from all available parameters.

In contrast to the work by both Davis *et al*. (1999) and Hendrikx *et al*. (2005), only natural avalanches were considered in this thesis, with artificially triggered avalanches excluded from both avalanche and non avalanche days. Davies *et al*. (1999) however states that the majority of avalanches are artificially triggered at the Alta and Mammoth study areas due to their rigorous avalanche control program. Additionally, Hendrikx *et al*. (2005) include all avalanches in their analysis, whether artificially triggered or naturally released as these form the current avalanche regime along the Milford Road, this equates to approximately equal numbers of each over the 17-year period of investigation. This discrepancy between data used in this thesis and published work may be one reason for the differences seen regarding the significance in the role played by wind drift factors by Davis *et al*. (1999) and Hendrikx *et al*. (2005) compared with the work in this thesis.

Another possibility is that the Fonnbu site may not be considered a particularly suitable place for wind measurements to be taken. This is due to it being at a valley site at some distance from the avalanche starting zones, and therefore not giving representative wind conditions for these zones.

5.3.3 Effectiveness in terms of avalanche day prediction

Two methods have been analysed for their application to correctly predict avalanche days with the use of the combined wind drift parameters. These are discussed in further detail in the following sections.

5.3.3.1 Probability plots

By looking at the probability plots in figures 4.6(a) and 4.6(b) created for two of the combined wind drift parameters from the Fonnbu data set and the corresponding parameters in the gridded data set, it is possible to predict, to a certain degree, the probability of an avalanche occurring. Using the maximum wind speed data to create the values in figure 4.6(b) leads to a much greater spread of wind drift values at specific probabilities. However this plot provides a better match between the gridded data and the Fonnbu data.

In comparison to the probability plots that have been created with just the precipitation parameters, the wind drift plots are obviously more complex to interpret due to the reason outlined in part 5.3.1. By increasing the complexity of the plot to include more parameters it would be expected that accuracy may also increase, however in this case there is greater potential for errors incurred from interpreting the wind drift factor. Probability plots, if considered individually, may therefore be inappropriate prediction methods if only due to the variation introduced on interpretation of the wind drift values.

5.3.3.2 Classification trees

As has been described in section 2.2.2.1 the temperature dependent wind drift parameter across the three day period created by Hendrikx *et al*. (2005) is considered the most effective parameter for avalanche day prediction, this is due to its placement as the initial splitting factor in the classification tree created. This is not the general case for the classification trees created in this thesis as only 16 % of the 50 trees use a wind drift parameter in the first split. This percentage decreases to approximately 10 % for the second place and third place splits. On accounting for the whole tree, approximately 40 % of the pruned trees created actually use one or more of the thirteen wind drift parameters created, 32 % of which are placed within the top three splits. This is to some extent more inline with the results found by Davis *et al*. (1999) in which, when using all parameters, precipitation factors were often ranked highest with snow drift factors within the top five splits but often not achieving the top splitting position. Contrary to Davis *et al*. (1999) however, is that *all* their tests indicated wind drift ranked in the top five, whereas almost 60 % of the 50 trees created in this thesis do not include any of the combined wind drift factors. One discrepancy here is that by setting the complexity parameter at 0.03, many of the trees created here only provide three splits. It is therefore unknown if any of the wind drift factors appear in the following two splits.

A comparison of wind drift occurrence within classification trees can also be made with Kronholm *et al*. (2006b). In their paper, although the majority of analysis undertaken discounted the use of the combined wind drift factor, on the occasion when they were used to create single element trees (using only the top split) these parameters featured highly. This is in contrast to results in this thesis. One reason for this may relate back to the difference between the data sets in tables 4.2 and 4.3, with the gridded data set used by Kronholm *et al*. (2006b) having much greater variation than the observed Fonnbu data used in this thesis for classification trees. This may account for the greater difference in wind drift values between avalanche and non avalanche days. Additionally, as stated in section 5.3.2.2 the Fonnbu site may not be the most suitable place to measure wind speed for use in the combined wind drift factors.

This difference in results of classification trees created by Davis *et al*. (1999), Hendrikx *et al*. (2005) and now in this thesis may be explained simply due to the fact that, at being at different locations, and with differing weather conditions, avalanche formation is influenced by different meteorological factors. Alternatively, the differences may be more feasible due to human variation in the weather sampling techniques, and/or the methods of statistical analysis. It has already been mentioned (section 3.4.4.1) that classification trees show great variability and instability depending on their growing methods and differences in random data sets used. In the case of Hendrikx *et al*. (2005), two trees were created and suggested for use in forecasting, it is however unknown if these were the only trees created or if this was one of the most commonly occurring tree formations from different sets of random data as is the case in this thesis.

5.4 Viability of Classification trees

5.4.1 Prediction of triggering parameters

All 45 parameters within the Fonnbu data set were initially implemented in growing the classification trees, this was despite some parameters showing no significant difference between avalanche days and non avalanche days in the Kruskal-Wallis test. All parameters were put forward for two main reasons; the first being that the Kruskal-Wallis test used only one set of random non avalanche days, whereas the classification tree analysis used 50 different sets of random non avalanche days. It is possible that these other random sets of data may have shown significant difference between avalanche days and non avalanche days where the set used in the K-W test failed to show a difference. Parameters were therefore not excluded due to one set of results showing insignificant difference. The second reason all parameters were applied in the classification trees is that due to the nature of the classification tree procedure, individual parameters can be influenced by each other and in certain combinations some parameters may have greater importance than if regarded individually. This is likely to be the case for the parameters TAM1day and FXX0day, both these are seen as important splitting factors in the first three splits of the classification trees. However, according to the K-W test, individually, these do not show any great significant difference between avalanche and non avalanche days.

By examining the top three splits of the classification trees created it was possible to identify some of the important avalanche triggering mechanisms. A surprising result appeared from table 4.7 showing TAX0day as the most frequently occurring factor within the top three splits, this is despite the exclusion of wet avalanche days. This result is in contrast to results by Davis *et al*. (1999) who indicate that temperature parameters rank lower than the wind, precipitation and combined wind drift parameters implemented. However a possible discrepancy with this statement is that when looking at the top three splits individually in this thesis, TAX0day does not enter into split one at all, but becomes much more apparent in split three. In comparison with Hendrikx *et al*. (2005), although only one full classification tree is presented it will be assumed that this is a typically expected representation; in this case, temperature parameters do not rank within the top three splits and only become apparent lower down. This is again in contrast to many of the trees created in this thesis where TAX0day ranks frequently within the top three splits.

There are several possible explanations for the high frequency of TAX0day as a significant splitting factor in this thesis, the first being that there are greater errors than expected in the initial observation and recording of the avalanche type. Within the study area, avalanches could be observed from a distance of up to 1 or 2 km away and at this distance it is obviously very difficult and also particularly subjective as to whether an avalanche is classed as wet, dry or of unknown type. An alternative suggestion is that, as this parameter relates to the maximum temperature recorded on the day, this provides an indication of large fluctuations in temperature. This is particularly apparent as the average threshold value is positive and given ‘greater than’ status, i.e. dry avalanches are expected under conditions when TAX0day is greater than about 4.5˚C and hence it follows that no avalanches are predicted with TAX0day less than 4.5˚C. To note here, is the temperature parameter highlighted by Hendrikx *et al*. (2005), although this parameter is ranked below the top three splits in the tree created, the parameter selected is the maximum temperature over the three day period. This relates to the principle outlined above, that the triggering of dry avalanches relates to large fluctuations in temperature in the run up to avalanche occurrence.

In relation to the precipitation parameters it is possible to see that they are by far the most predominant of factors in the first place split with RR2day and RR4day occurring most frequently closely followed by RR5day and RR3day. Both these latter parameters are important factors present in previous literature from the study area. Bakkehøi (1987) presented RR3day as a good predictive measure for avalanche triggering. Additionally, RR5day is a predominant feature in the first splitting position of classification trees created by Kronholm *et al*. (2006a; 2006b). Although the predominance of precipitation is of little surprise, it was expected that in this thesis the combined wind drift parameters would feature more highly within the classification trees as they have in previous literature (Hendrikx *et al*., 2005; Davis *et al*., 1999). In this thesis however, when accounting for the top three splits as a whole, wind drift factors appear least frequently coming after various precipitation, temperature, and wind parameters. One point of note, however, is that although precipitation is predominant within the first splitting position, on the occasions when a factor of precipitation is not chosen here, then one of the wind drift factors is always chosen instead. On these occasions this is significant for comparison with Hendrikx *et al*. (2005) and Davis *et al*. (1999), however it also implies that precipitation still has a large influence as an important triggering mechanism for dry avalanche creation.

5.4.2 Implementation of threshold values

Looking in detail at table 4.6 it is possible to see relatively large variations in threshold values, for instance, in the first split alone, RR2day has threshold values varying from approximately 7 to 25 mm. RR4day has a similar spread for threshold values between about 14 to 35 mm within the first splitting position. Although averages have been calculated for the most frequently occurring parameters in the top three splits in table 4.7, a margin of freedom must be allowed for when viewing these averages, due to the difficulty encountered with calculation. This was necessary as values are defined as greater than or less than, rather than equal to a certain number. In addition, threshold values are likely to have some dependence on the other parameters in the surrounding splitting positions therefore average values may be suggested more as a measure of scale for each factor rather than as an actual threshold for prediction.

Even with comparable trees when parameters chosen for each splitting position are the same, there is still variation of the threshold parameters. A good example is seen in trees 40 and 11 (see Appendix 4), both of which use RR4day (split 1), FXX0day (split 2) and TAX0day (split 3). Threshold values for these parameters vary between almost 1 mm for RR4day, although this may not be much, FXX0day varies by 8 m/s and TAX0day varies by 1.35˚C. These latter two variations are quite large considering the ultimate spread of data values within each parameter. The reason for this variation is down to the different set of random data used for the non avalanche days. This fact must be taken into consideration with the application of thresholds, which will invariably have different values due to the inherent variability of current and preceding weather conditions.

In terms of the wind drift parameters, it has been stated in previous sections (particularly 5.3.1) that threshold values are difficult to interpret. Although this is the case, on inspecting parameter RRFXX1 more closely, threshold values are relatively similar across the trees in which this parameter is present. There is only small variation in relative terms from the average of 200,000 mm(m/s)^4 presented in table 4.7. This average can also be cross referenced with the probability plot of this factor in figure 4.6(b). Here a wind drift factor of 200,000 mm(m/s)^4 correlates to a 50 % probability of avalanche occurrence based on the corresponding Fonnbu data.

A final point to note which has been briefly mentioned in section 5.4.1 is the threshold values for temperature, particularly that of TAX0day. Dry avalanches are generally predicted when maximum temperatures exceed a threshold value of around plus 4˚C, as seen in table 4.6. In contrast, the classification tree applied by Hendrikx *et al.* (2005) shows a threshold of < 2.9˚C, implying that dry avalanches occur when maximum temperatures over the 72 hour period (in this case) do not exceed 2.9˚C. This difference is difficult to explain, however, one reason could be due to the large elevation difference and hence temperature difference between the Fonnbu weather station and avalanche release zones. Another explanation may be that in this thesis temperature has only been provided for the day of the avalanche and the preceding day. Over this short time period larger temperature fluctuations are required to create a fracture zone triggering avalanches. This is opposed to over a slightly longer period of three days, used in analysis by Hendrikx *et al*. (2005), when smaller fluctuations closer to zero degrees are enough to cause an avalanche. This has been presented in section 2.2.4 as McClung and Schaerer (1993) and SLF (2006) state that temperature fluctuations cause the snow surface to melt and refreeze. Over a three day period this surface layer can become buried beneath new snowfall becoming a subsequent sliding layer.

5.4.3 Applicability to this site and others

In general it has been seen that other studies show wind drift parameters as a stronger avalanche predictor than just precipitation alone (Hendrikx *et al*., 2005; Davis *et al*., 1999). The result in this thesis that precipitation factors alone rank highest among classification tree splits is therefore somewhat surprising. What this report does highlight is that within the study area in Stryn, summed precipitation over the two to four day period is the predominant splitting factor in initial classification. This is supported by the results in the same area to predict avalanche occurrence based on the probability measured across the three day sum by Bakkehøi (1987) and the importance of precipitation and the five day sum by Kronholm *et al*. (2006a; 2006b). One particular reason why wind parameters were not seen to play such a large role at this site compared to other areas (Davis *et al*., 1999; Hendrikx *et al*., 2005) could be due to the location of the weather station in the valley. The wind speed measured at Fonnbu, being at a valley site, may therefore not be a realistic measure of the wind speeds experienced at higher altitudes where the majority of the avalanche starting zones lie. An alternate weather station at Kvitenova within the study area is at a hill top site, this is therefore more exposed, and probably more representative of the wind speeds experienced around avalanche starting zones. Data from this weather station was however unavailable at the time of writing this thesis.

The creation of classification trees in this thesis has involved looking in detail at the script used with the R program. It has become obvious that there is no standard method for creating classification trees, there are instead a number of variations and adjustments that can be made to any one tree. These vary not only upon details within the script such as complexity and amount of pruning of the tree but also on which tree packages are used in conjunction with the R program. Additionally, even on setting all the variables within the script to a constant, the creation of 50 different classification trees with varying non avalanche days highlights the random and often unstable nature of the classification tree procedure. Additionally, in this thesis it was apparent that the accuracy of the trees fluctuates with misclassification rates of between 20 % and 30 %. For forecasting purposes these are considered too high by Föhn (1998) in Kronholm *et al*. (2006a) who states that misclassification rates should ideally be less than 20 %.

The above paragraph gives evidence that classification trees give very variable results not only when applied to different study sites and data, but also when used at just one location with a partially variable data set. One point to mention is, however, that classification trees may enhance some prediction methods as they have the ability to show probability of an avalanche occurring and also the probability of no avalanche occurring. This is therefore more superior to probability plots which can only indicate the probability of an avalanche occurring.

Finally, it may be important to mention the findings of this thesis with regards to the influence from broader climatic factors. This includes both long term cycles such as the NAO, in addition to anthropogenic greenhouse warming effects. Both a positive NAO and future predictions for the influence of anthropogenic factors indicate that warmer and wetter weather are likely to be experienced in western Norway and across much of northern Europe (Keylock, 2003; Ulbrick and Christoph, 1999; Hanssen-Bauer and Førland, 2000). This must therefore be noted in terms of the implication this may have on the classification trees created. From the trees in this thesis it has been shown that both precipitation and temperature related parameters are important in terms of distinguishing between dry avalanche and non avalanche days. The increase of these, influenced by predicted climate change is likely to affect the forecasting ability of the classification trees. Although, as the trees currently suggest, an increase in temperature and precipitation is associated with an increase in the number of avalanche days. However, much further investigation of the classification trees, in relation to future climate change, is required in order to analyse this relationship effectively before such a notion could be accepted.

Chapter 6: Conclusion and Recommendations

6.1 Overall summary

Snow avalanches are a common occurrence throughout the winter months in many areas of the world. This thesis has looked in detail at the area surrounding Grasdalen in western Norway, in order to establish certain relationships between weather parameters and avalanche occurrence. Two main sets of weather data; an extrapolated gridded set and an observed set from the Fonnbu weather station have been analysed using varying statistical procedures.

Initially, exploratory analysis of both data sets was carried out. Results indicated measures of precipitation to be approximately three times greater in the gridded data compared to the Fonnbu data. In terms of wind speed; the majority of parameters show values consistent with those required to create substantial wind drift. The gridded data set showed higher average values of wind speed, but lower maximum values of wind speed in comparison to the Fonnbu data set. Other preliminary results reveal a predominant south-westerly wind direction on avalanche days. Finally a number of wind drift parameters were calculated for both data sets. These were based on the definition by Davis *et al*. (1999) as the product of precipitation and wind speed to the fourth power.

Statistical procedures included the creation of a number of cumulative probability plots. These have outlined probabilities of avalanche occurrence for the ten most frequent avalanche paths from the gridded data set in addition to using avalanche day data from the Fonnbu data set. Comparisons were made between these and similar plots created by Bakkehøi (1987). Results showed that for certain precipitation levels, the Fonnbu data set showed consistently higher probabilities for avalanche occurrence. Unfortunately this data set did not distinguish between individual paths. Closer comparison of the gridded data probability plots and similar plots by Bakkehøi (1987) were made for some individual paths. In three out of the four paths analysed, good correlation was seen at the 50 % probability level. For these, approximately 50 mm of precipitation was required at this probability level. However, beyond about 80 % probability of avalanche occurrence, there was little correlation between the plots by Bakkehøi (1987) and the gridded data plots. To conclude, probability plots can provide basic threshold values for certain precipitation parameters, however these can vary greatly for differing avalanche paths.

Fifty classification trees were created using only the Fonnbu data set to distinguish between dry avalanche and non avalanche days. This data set had not been used previously for such a procedure. The classification trees have provided easily interpretable results, incorporating multiple parameters, with the most important splitting parameters recorded within the top three positions of the trees. Results were surprising as the combined wind drift factors were not portrayed as important splitting factors. This contradicts the majority of results seen in previous literature (Hendrikx *et al.*, 2005; Davis *et al.*, 1999; Kronholm *et al.*, 2006b). One significant reason for this may be that the Fonnbu weather station, being located in a valley site, is not in the most ideal position to provide weather conditions representative of those experienced in the avalanche starting zones.

An important splitting factor highlighted by the classification trees is precipitation, particularly over the two and four day sums. This is shown by their predominance in the top splitting position. Following this, splits two and three show the significance of temperature particularly TAX0day (maximum temperature on the preceding day). The high threshold values for this parameter indicate that large fluctuations in temperature over a short time period are important to distinguish dry avalanche days from non avalanche days. In relation to the combined wind drift parameters, this was the first time these were created for the Fonnbu data set and applied to classification trees. These parameters feature to some extent (32 % of the time) within the top three splits of the classification trees. Of these, the wind drift derived using maximum wind speeds on the day of the avalanche occur most frequently. One disadvantage apparent for the combined wind drift parameter relates to the complex interpretation of threshold values.

In conclusion, classification trees can generally be considered to provide a good tool for prediction, by distinguishing between avalanche and non avalanche days. However, in this thesis, the main purpose of these classification trees was to show the most important splitting parameters and associated threshold values. Many were therefore pruned too harshly for prediction purposes. This is apparent as misclassification rates are greater than the 20 % that would generally be accepted for forecasting (Föhn, 1998 in Kronholm *et al*., 2006a).

6.2 Recommendations for further work

This thesis has provided classification analysis for the first time on the Fonnbu data set, with classification trees distinguishing between dry naturally occurring avalanche days and non avalanche days. Further work with this data set could provide broader analysis of classification trees, using all avalanche types, and inclusion of artificially triggered avalanches as has been the case in previous literature (Hendrikx *et al*., 2005; Davis *et al*., 1999). Additionally, the study could be extended regarding the combined wind drift parameters. This could include the creation of a temperature-combined wind drift factor, particularly as temperature alone was considered an important factor in splits two and three of the classification trees created in this thesis.

Alternatively, data from the weather station at Kvitenova could be applied with similar statistical procedures as have been carried out in this thesis. Being a hill top site this may better represent weather conditions within the nearby avalanche starting zones. Classification trees created with this data may therefore provide a better rate of occurrence of combined wind drift factors, more comparable with results by Hendrikx *et al*. (2005) and Davis *et al*. (1999). Wind directional data from this site at Kvitenova could also be included in some wind drift parameters to examine the importance of wind direction for avalanche day prediction.

Finally, it could be interesting to explore ways to make the classification tree procedure user friendly in terms of a warning / forecasting method. This could be in terms of a remotely operated program using the relevant values from the data recorded at the weather station.

Reference List

Akitaya, E. (1974) Studies on depth hoar [In] *Snow mechanics Symposium, International Association of Hydrological Sciences* **114** pp42-48.

Akkouratov, V. N. (1965) Meteorological conditions of avalanche formation in the Khibiny [In] *International Symposium on Scientific Aspects of Snow and Ice Avalanches*, Davos pp 35-42.

Ambach, W. and Howorka, F. (1965) Avalanche activity and free water content of snow at Obergurgl (1980m a.s.l., Spring 1962) [In] *International Symposium on Scientific Aspects of Snow and Ice Avalanches*, Davos pp65-72.

Bakkehøi, S. (1987) Snow avalanche prediction using a probabilistic method. [In] *Avalanche formation, movements and effects; Proceedings of the Davos Symposium, International Association of Hydrological Sciences* **162** pp549-555.

Barry, R. G. (1981) *Mountain weather and climate*. Methuen, New York. pp313.

Blikra, L. H. and Nemec, W. (2000) Reply to: Facies and palaeoclimatic significance of post-glacial colluvial deposits. *Sedimentology* **47** pp1058-1068.

Birkeland, K., Johnson, R. and Schmidt, S. (1996) Near-surface faceted crystals: Conditions necessary for growth and contribution to avalanche formation, Southwest Montana, U.S.A. [on-line] available from: http://www.mtavalanche.com/articles/art2.htm [Accessed 04.07.06].

Birkeland, K.W. (2001) Spatial patterns of snow stability throughout a small mountain range. *Journal of Glaciology* **47**(157) pp176-186.

Bois, P., Obled, C. and Good W. (1974) Multivariate data analysis as a tool for day-by-day avalanche forecast. [In] *Snow Mechanics; Proceedings of the Grindelwald Symposium, International Association of Hydrological Sciences* **114** pp391-403.

Brabec, B. and Meister, R. (2001) A nearest-neighbor model for regional avalanche forecasting. *Annals of Glaciology* **32** pp130-134.

Breiman, L., Freidman, J., Olshen, R. and Stone, C. (1993) *Classification and regression trees*. Chapman and Hall, New York. pp358.

Budd, W. R., Dingle, R. and Radok, U. (1966) The Byrd snowdrift project: Outline and basic results. *Studies in Antarctic Meteorology, American Geophysical Union, Antarctic Research Service* **9** pp71-134.

Butler, D. R. (1986) Snow-avalanche hazards in Glacier National Park, Montana: Meteorologic and climatologic aspects. *Physical Geography* **7**(1) pp72-87.

Custer, S. (2005) Snowfall intensity criteria for avalanche [on-line]. Available from: http://www.homepage.montana.edu/~uessc/esci450/SnowIntensity.htm [Accessed 30/11/06].

Davis, R.E., Elder, K., Howlett, D. and Bouzaglou, E. (1999) Relating storm and weather factors to dry slab avalanche activity at Alta, Utah, and Mammoth Mountain, California, using classification and regression trees. *Cold Regions Science and Technology* **30** pp79-89.

de Quervain, M. R. (1965) Problems of avalanche research [In] *International Symposium on Scientific Aspects of Snow and Ice Avalanches*, Davos pp15-22.

Dyunin, A. K. and Kotlyakov, V. M. (1980) Redistribution of snow in mountains under the effect of heavy snow-storms. *Cold Regions Science and Technology* **3** pp287-294.

Fitzharris, B., Lawson, W. and Owens, I. (1999) Research on glaciers and snow in New Zealand. *Progress in Physical Geography* **23**(4) pp469-500.

Furdada, G., Martí, G., Oller, P., García, C., Mases, M. and Vilaplana, J.M. (1995) Avalanche mapping and related GIS applications in the Catalan Pyrenees. *Surveys in Geophysics* **16** pp681-693.

Föhn, P. M. B. (1998) An overview of avalanche forecasting models and methods [In] Hestnes, E. (Ed.) 25 years of snow avalanche research, Voss, Norway 12-16th May 1998. Norwegian Geotechnical Institute, Oslo, Norway pp19-27.

Gassner, M. and Brabec, B. (2002) Nearest neighbour models for local and regional avalanche forecasting. *Natural Hazards and Earth System Sciences* **2** pp247-253.

Graphpad (1999) Interpreting the Kruskal-Wallis test [on-line]. Available from: http://www.graphpad.com/articles/interpret/ANOVA/kruskal_wallis.htm [Accessed 03/05/07].

Gruber, U. and Margreth, S. (2001) Winter 1999: a valuable test of the avalanche-hazard mapping procedure in Switzerland. *Annals of Glaciology* **32** pp328-332.

Hanssen-Bauer, I. and Førland, E. (2000) Temperature and precipitation variations in Norway 1900-1994 and their links to atmospheric circulation. *International Journal of Climatology* **20** pp1693-1708.

Hendrikx, J., Owens, I., Carran, W. and Carren, A. (2005) Avalanche activity in an extreme maritime climate: The application of classification trees for forecasting. *Cold Regions Science and Technology* **43** pp104-116.

Jamieson, B. (2004) Between a slab and a hard layer [on-line]. Available from: http://www.schulich.ucalgary.ca/Civil/Avalanche/Papers/PBCrustsAn1.pdf [Accessed 04/12/06].

Keylock, C. J. (2003) The North Atlantic Oscillation and snow avalanching in Iceland. *Geophysical Research Letters* **30**(5) pp58-61.

Kistler, R. (2001) The NCEP-NCAR 50-Year Reanalysis: Monthly means CD-ROM and documentation. *Bulletin of the American Meteorological Society* **82**(2) pp247-267.

Kotlyakov, V. M. and Plam, M. Y. (1965) The influence of drifting on snow distribution in the mountains and its role in the formation of avalanches [In] *International Symposium on Scientific Aspects of Snow and Ice Avalanches*, Davos pp53-60.

Kristensen, K., Harbitz, C.B. and Harbitz, A. (2003) Road traffic and avalanches – Methods for risk evaluation and risk management. *Surveys in Geophysics* **24** pp603-616.

Kronholm, K., Vikhamar-Schuler, D., Jaedicke, C., Isaksen, K., Sorteberg, A. and Kristensen, K. (2006a) Forecasting snow avalanche days from meteorological data using classification trees; Grasdalen, Norway. *Proceedings of the International Snow Science Workshop* [on-line] available from: http://www.uib.no/People/gbsag/Kronholm_etal_2006.pdf [Accessed 18.02.07].

Kronholm, K., Vikhamar-Schuler, D., Jaedicke, C., Isaksen, K., Sorteberg, A., and Kristensen, K. (2006b) Meteorological triggering elements for snow avalanche release; Grasdalen, Western Norway. S*ubmitted to CRST special issue from the ISSW 2006* pp34.

LaChapelle, E. (1961) Snow Avalanches. *Agriculture Handbook* **194** U.S. Department of Agriculture, Forest Service.

LaChapelle, E. (1962) The chemical modification of depth hoar [In] Kingery, W. D. (Ed.) *Ice and snow properties, processes and applications*. M.I.T. Press, Massachusetts pp586-593.

Landry, C., Birkeland, K., Hansen, K., Borkowski, J., Brown, R. and Aspinall, R. (2002) Snow stability on uniform slopes: implications for avalanche forecasting. *Proceedings of the International Snow Science Workshop 2002.* [on-line] available from: http://www.avalanche.org/~issw2004/issw_previous/2002/flashsite/Snow_Stability/LANDRY%20ORAL.PDF [Accessed 18.02.07].

Lied, K., Instantes, B., Domås, U. and Harbitz, C. (1998) Snow avalanche at Bleie, Ullensvang, January 1994. *Norwegian Geotechnical Institute Publication* **203** pp175-181.

Logan, N. (1998) Forecasting avalanches from atmospheric cross-sections [on-line]. Available from: http://www.avalanche.org/~issw98/poster_sessionrisk_management.htm [Accessed 19/11/06].

Male, D. H. (1980) The seasonal snowcover [In] Colbeck, S. C. (Ed.) *Dynamics of snow and ice masses* Academic Press, New York pp 305-390.

Maggioni, M. and Gruber, U. (2002) The influence of topographic parameters on avalanche release dimension and frequency. [In] Stevens, J. (Ed.) *Proceedings of the International Snow Science Workshop 2002*, B.C., Canada pp97-103.

Mathworks (2007) Statistics toolbox kruskalwallis [on-line]. Available from: http://www.mathworks.com/access/helpdesk/help/toolbox/stats/index.html?/access/helpdesk/help/toolbox/stats/kruskalwallis.html&http://www.google.co.uk/search?hl=en&q=Kruskal-wallis&meta= [Accessed 04/04/07].

McClung, D. and Schaerer, P. (1993) *The avalanche handbook*. The Mountaineers, Washington. pp 271.

McElwaine, J., Hachikubo, A., Nemoto, M., Kaihara, T., Yamada, T. and Nishimura, K. (2000) Observations and simulations of the formation of the faceted snow crystals in the weak-layer of the 1998 Niseko Haru no Taki avalanche. *Cold Regions Science and Technology* **31**(3) pp235-247.

Met Office (2006) Statistical prediction of the winter NAO [on-line]. Available from: http://www.metoffice.gov.uk/research/seasonal/regional/nao/ [Accessed 05/12/06].

Norwegian Meteorological Institute (2004) Informasjon om snøkart [on-line]. Available from: http://met.no/snokart/informasjon.html [Accessed 23/05/07]

Norwegian Meteorological Institute (2006) The climate of Norway [on-line]. Available from: http://www.met.no/english/climate/index.html [Accessed 19/11/06].

Owens, I. F. and Weir, P. L. (1992) The Broken River avalanche of August 1992; Craigieburn Range, New Zealand [In] *Proceedings of the International Snow Science Workshop*, Colorado, Avalanche Information Centre pp335-341.

Perla, R. I. (1978) Failure of snow slopes [In] Voight, B. (Ed.) *Rockslides and Avalanches, 1. Natural Phenomena; developments in geotechnical engineering* **14A** Elsevier Science Publishing, Oxford pp731-752.

Perla, R. I. (1980) Avalanche release, motion, and impact [In] Colbeck, S. C. (Ed.) *Dynamics of Snow and Ice Masses*, Academic Press, New York pp397-462.

Pomeroy, J. W., Gray, D. M. and Landline, P. G. (1991) Modelling the transport and sublimation of blowing snow on the prairies. [In] *Proceedings of the 48th Eastern Snow Conference* pp175-188.

Pomeroy, J. W. and Gray, D. M. (1995) *Snowcover; accumulation, relocation and management.* National Hydrology Research Institute, Canada. pp144.

Pomeroy, J. W., Gray, D. M., Shook, K. R., Toth, B., Essery, R. L. H., Pietroniro, A. and Hedstrom, N. (1998) An evaluation of snow accumulation and ablation processes for land surface modelling. *Hydrological Processes* **12** pp2339-2367.

R Development Core Team (2005) R: A language and environment for statistical computing, version 2.2.1. R Foundation for Statistical Computing, Vienna, Austria [on-line]. Available from: http://www.R-project.org [Accessed 2007].

RTD (2006) Management of natural hazards. RTD Info 24 Magazine for European Research [on-line]. Available from: http://ec.europa.eu/research/rtdinfo/en/24/07.html [Accessed 04.07.06].

SAIS (1999) Forecast archive, Scotland Avalanche Information Service [on-line]. Available from: http://www.sais.gov.uk/forecast_archive/99-00/1999.12.24 [Accessed 21/11/06].

Schweizer, J., Jamieson, B. and Schneebeli, M. (2003) Snow avalanche formation. *Reviews of Geophysics* **41**(4) pp1016-1039.

SLF (2006) Glossary snow and avalanches; Swiss Federal Institute for Snow Avalanche Research, European avalanche forecasting services [on-line]. Available from: http://wa.slf.ch/index.php?id=278#723 [Accessed 30/11/06].

StatSoft (2003) Classification Trees [on-line]. Available from: http://www.statsoft.com/textbook/stclatre.html [Accessed 16/03/07].

Tabler, R. D., Pomeroy, J. W. and Santana, B. W. (1990) Drifting snow. [In] Ryan, W. L. and Crissman (Eds.) *Cold Regions Hydrology and Hydraulics*, American Society of Civil Engineers, New York pp95-146.

Takeuchi, M. (1980) Vertical profiles and horizontal increase of drift snow transport. *Journal of Glaciology* **26**(94) pp481-492.

Therneau, T. M. and Atkinson, B. (2005) rpart: Recursive partitioning, R package version 3.1-27 [on-line]. Available from: http://mayoresearch.mayo.edu/mayo/research/biostat/splusfunctions.cfm [Accessed 24/05/07].

Ulbrich, U. and Christoph, M. (1999) A shift in the NAO and increasing storm track activity over Europe due to anthropogenic greenhouse gas forcing. *Climate Dynamics* **15** pp551-559.

UNESCO (1981) *Avalanche atlas, Illustrated international avalanche classification.* United Nations Educational, Scientific and Cultural Organisation, Switzerland. pp265.

Walsh, S.J., Butler, D.R., Brown, D.G. and Bian, L. (1990) Cartographic modeling of snow avalanche path location within Glacier National Park, Montana. *Photogrammetric Engineering and Remote Sensing* **56**(5) pp615-621.

Weka (2007) Classification Trees, Waikato Environment for Knowledge Analysis, University of Weikato [on-line]. Available from http://wekadocs.com/node/2 [Accessed 17/03/07].

Wheeler, D., Shaw, G. and Barr, S. (2004) *Statistical Techniques in Geographical Analysis* (3rd edition). David Fulton, London pp342.

Wikipedia (2007a) Stryn map, Sogn and Fjordane [on-line]. Available from: http://en.wikipedia.org/wiki/Sogn_and_Fjordane [Accessed 04/03/07].

Wikipedia (2007b) Decision Tree [on-line]. Available from: http://en.wikipedia.org/wiki/Decision_tree [Accessed 26/05/07].

Wikipedia (2007c) Kruskal-Wallis one-way analysis of variance [on-line]. Available from: http://en.wikipedia.org/wiki/Kruskal-Wallis_one-way_analysis_of_variance [Accessed 03/05/07].

Zingg, T. (1965) Relation between weather situation, snow metamorphism and avalanche activity [In] *International Symposium on Scientific Aspects of Snow and Ice Avalanches*, Davos pp61-64.

Appendix 1: Avalanche name and code

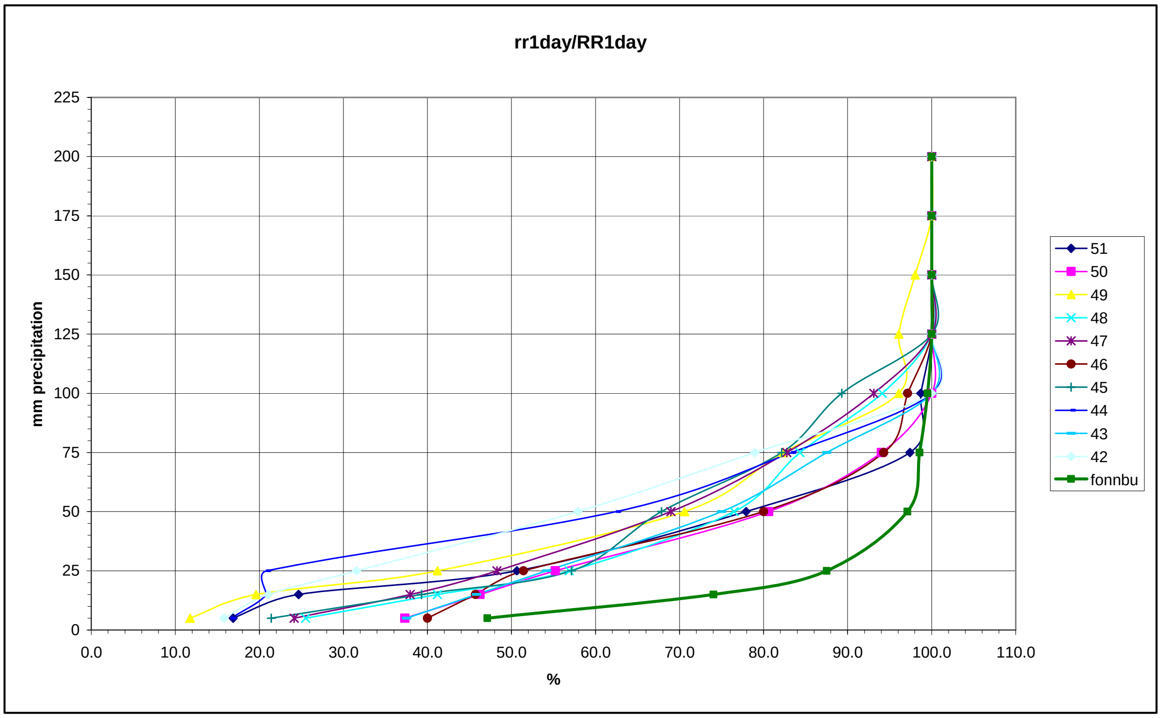
The top ten most frequently occurring avalanches are indicated in bold and the predominant aspect for each avalanche path is provided for reference.

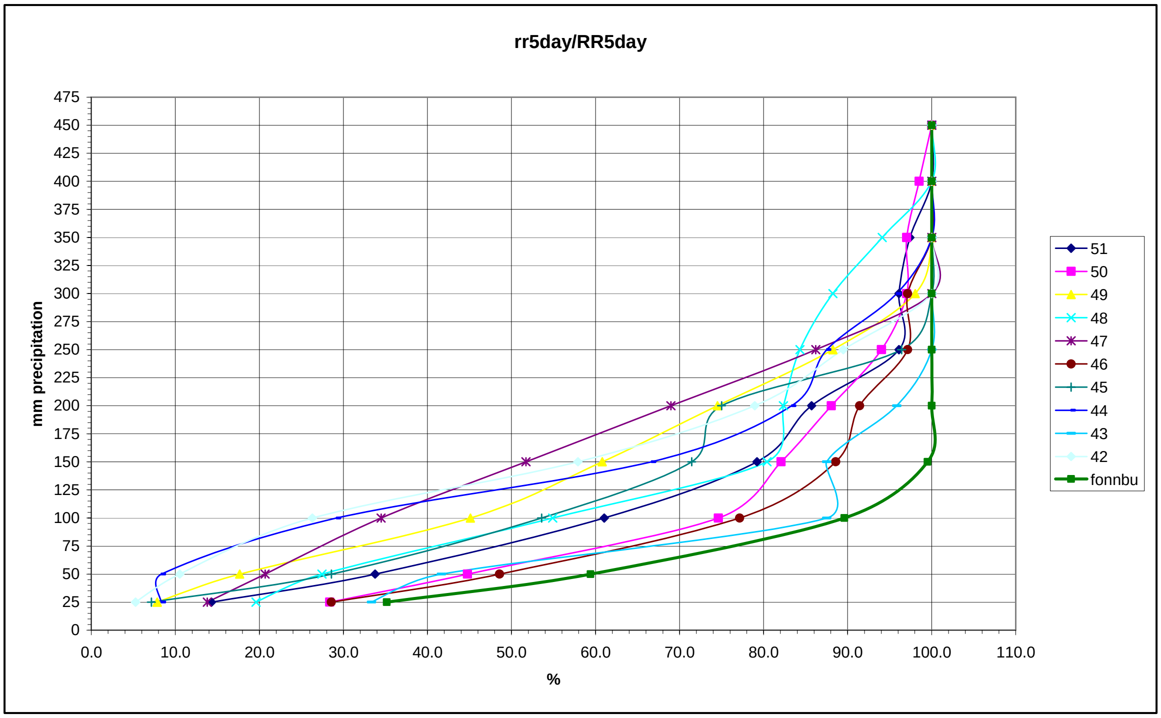
Slide_PlaceName	**Code name**	**Global_ExpositionCode**
Sætreskarsfjellet / Lifonn	**51**	**E**
Storfonn	**50**	**SW**
Ryggfonn	**49**	**NW**
Svartefjellet	**48**	**W**
Breiskredfonn	**47**	**SE**
Raffelsteinfonn	**46**	**SE**
Storurfonn	**45**	**S**
Napefonn	**44**	**NW**
Raudnova N	**43**	**NW**
Grasdalstunnel W	**42**	**SW**
Raudnova NW	41	NW
Godtidfonn	40	NW
Blåbærfonn	39	NW
Grasdalstunnel E	38	S
Fonnbu, NE for	37	SW
Ryggfonn, W for	36	NW
Svartebardskaret	35	NW
Sætreskarsfjellet topp	34	NE
Raudnova topp NE	33	NE
Blåbærfonn, W for	32	NW
Fonnbu, E for	31	NW
Långesvingfonn	30	NW
Skjæringsdalen S	29	E
Råsdalsfonn	28	SW
Resfonn	27	S
Grasdalsvatnet, W for	26	E
Grasdalsvatnet, SW for	25	NE
Raudnova topp N	24	N
Ospelitunnel N	23	W

Midtfonn	22	NW
Blåbærfonn, E for	21	NW
Sætreskaret	20	SW
Raudnova W	19	NW
Ospelitunnel S	18	S
Oppljostunnel W	17	W
Oppljostunnel E	16	NA
Grasdalsvatnet, NE for	15	SW
Breifonn Skåre	14	NW
Stavbrekka	13	SW
Skåre	12	S
Grasdalsbreen S	11	E
Fonnbu, W for	10	SE
Breifonn Ospeli	9	W
Oppljosegga SE	8	SE
Sætreskarsvatnet, NW for	7	SE
Skjæringsdalen Mitt	6	E
Skjæringsdalssætra bru	5	NW
Storefonnhyrna	4	E
Svartebardskaret topp	3	NW
Oppljosegga N	2	N
Fonnbu	1	NW

Appendix 2: Probability plots

These probability plots are for the one and five day sums of precipitation for the gridded and Fonnbu data with summarised results in table 4.4.





Appendix 3: R script

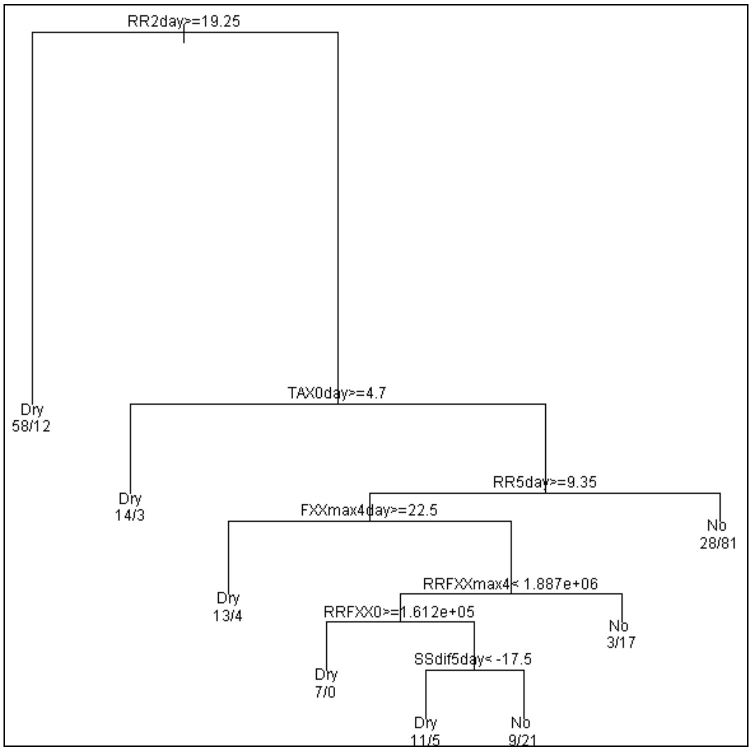
```
########################################################################
##############################
#
# NAME
#     classificationtrees_dryonly.R
#
# PURPOSE
#     - Load data from met dataset
#     - Select dry snow avalanche days
#     - Build full classification tree
#     - Prune classification tree based on 10-fold cross validation
#
#
########################################################################
##############################

#---------------------
#----- LOAD DATA -----
#---------------------

data.all <- read.table("2aval_met_mod.txt", header=TRUE)
summary(data.all)

#-----------------------
#----- SELECT DATA -----
#-----------------------

# EITHER select data with some missing met. parameters
data.dry <- subset(data.all, (n.metNAs<20 & nAval.dry>0))
# OR select dry avalanches with all met info
#data.dry <- subset(data.all, (allmet==TRUE & nAval.dry>0))
type <- rep("Dry", length(data.dry[,1]))
data.dry <- data.frame(data.dry, type)

# EITHER all no-avalanche days with all met info
data.noaval <- subset(data.all, (allmet==TRUE & nAval==0))
# OR no-avalanche days with some met. parameters missing
#data.noaval <- subset(data.all, (n.metNAs<10 & nAval==0))
type <- rep("No", length(data.noaval[,1]))
data.noaval <- data.frame(data.noaval, type)

# find out how many observations there are in the AVAL dataset and
select as many from the NOAVAL
n.aval.obs <- length(data.dry[,1])
print(n.aval.obs)
n.noaval.obs <- length(data.noaval[,1])
print(n.noaval.obs)

n.random.select <- 0
if (exists("random.select")) {rm(random.select)}
while (n.random.select < n.aval.obs) {
      # find the number of random numbers to generate
      n.generate <- n.aval.obs - n.random.select
```

```
      # generate the random numbers rounded
      new.random.select <- round(runif(n.generate, min=0,
max=n.noaval.obs), digits=0)
      # add the data to the existing dataset
      if (exists("random.select"))
            {random.select <- c(random.select, new.random.select)} else
            {random.select <- new.random.select}
      # find all the unique values and count
      random.select <- unique(random.select)
      n.random.select <- length(random.select)
}

#random.select

#write.table(random.select)

# finally selct the data
selected <- vector(mode="logical", length=n.noaval.obs)
selected[random.select] <- TRUE
data.noaval.select <- subset(data.noaval, (selected))

# add the two tables together for easier use
data.use <- rbind(data.dry, data.noaval.select)

#-----------------------
#----- RPART TREES -----
#-----------------------

# load the module
library(rpart)

# build the full rpart tree
rpart.dry.full <- rpart(type~RR0day
      +RR1day
      +RR2day
      +RR3day
      +RR4day
      +RR5day
      +TAM1day+TAM0day+TAX0day+TAX1day+TAN0day+TAN1day
      +SS0day+SS1day+SSdif1day+SSdif2day+SSdif3day+SSdif4day+SSdif5day
      +FFM0day+FFM1day+FFX0day+FFX1day
      +FXM0day+FXM1day+FXXmax1day+FXXmax2day+FXXmax3day+FXXmax4day+FXXm
ax5day
      +FXX0day+FXX1day
      +RRFFM0+RRFFM1+RRFFX0+RRFFX1+RRFXM0+RRFXM1+RRFXX0+RRFXX1
      +RRFXXmax1+RRFXXmax2+RRFXXmax3+RRFXXmax4+RRFXXmax5
      ,data=data.use,
      method="class"
)

# plot the complexity parameter for the full tree
plotcp(rpart.dry.full)

# set the complexity parameter that seems most suitable.
# based on a few runs on the data a CP of around 0.04 to 0.02 seems
most suitable.
# setting a higher CP means smaller tree and vice versa
```

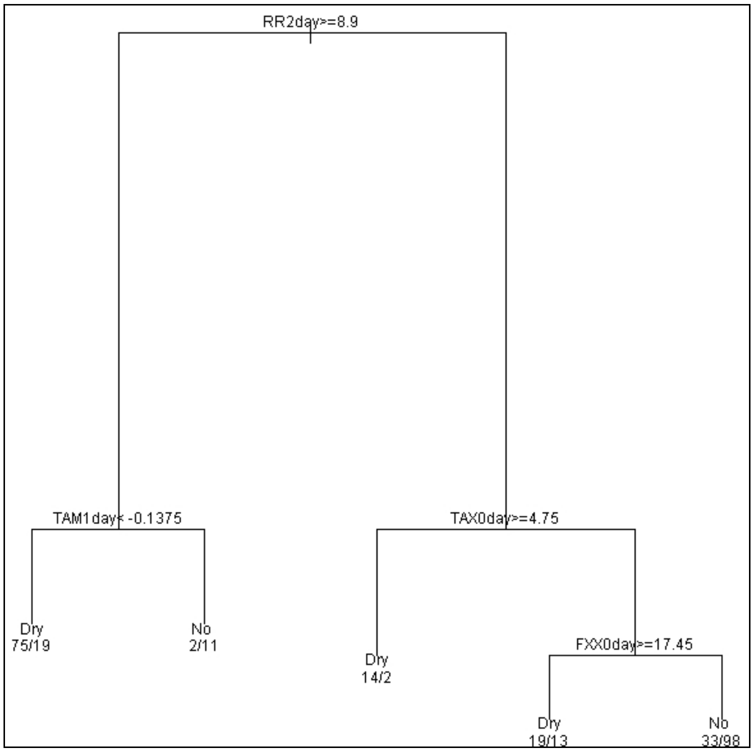
```
set.cp <- 0.03

# rebuild the tree based to the optimal size based on the complexity
parameter
rpart.dry.best <- rpart(type~RR0day
      +RR1day
      +RR2day
      +RR3day
      +RR4day
      +RR5day
      +TAM1day+TAM0day+TAX0day+TAX1day+TAN0day+TAN1day
      +SS0day+SS1day+SSdif1day+SSdif2day+SSdif3day+SSdif4day+SSdif5day
      +FFM0day+FFM1day+FFX0day+FFX1day
      +FXM0day+FXM1day+FXXmax1day+FXXmax2day+FXXmax3day+FXXmax4day+FXXm
ax5day
      +FXX0day+FXX1day
      +RRFFM0+RRFFM1+RRFFX0+RRFFX1+RRFXM0+RRFXM1+RRFXX0+RRFXX1
      +RRFXXmax1+RRFXXmax2+RRFXXmax3+RRFXXmax4+RRFXXmax5
      ,data=data.use,
      method="class", # builds a classification tree
      control=rpart.control(cp=set.cp)
)

# print the tree
print(rpart.dry.full)

# plot the tree
par(mar=c(0,0,0,0))
plot(rpart.dry.best); text(rpart.dry.best, cex=0.75, use.n=TRUE)
summary(rpart.dry.best)
```

Appendix 4: Classification tree diagrams

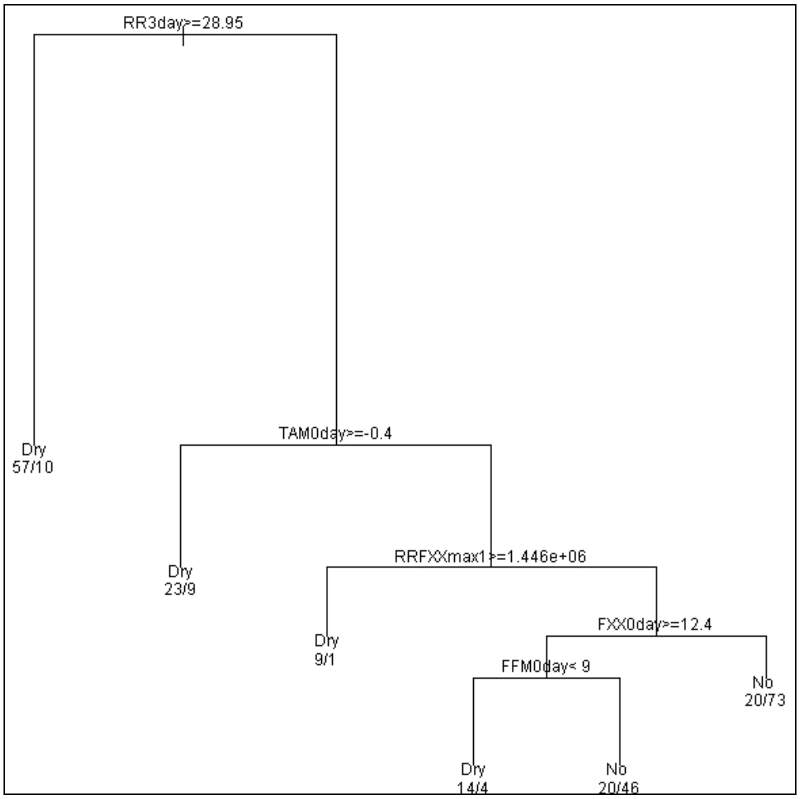
RUN 1



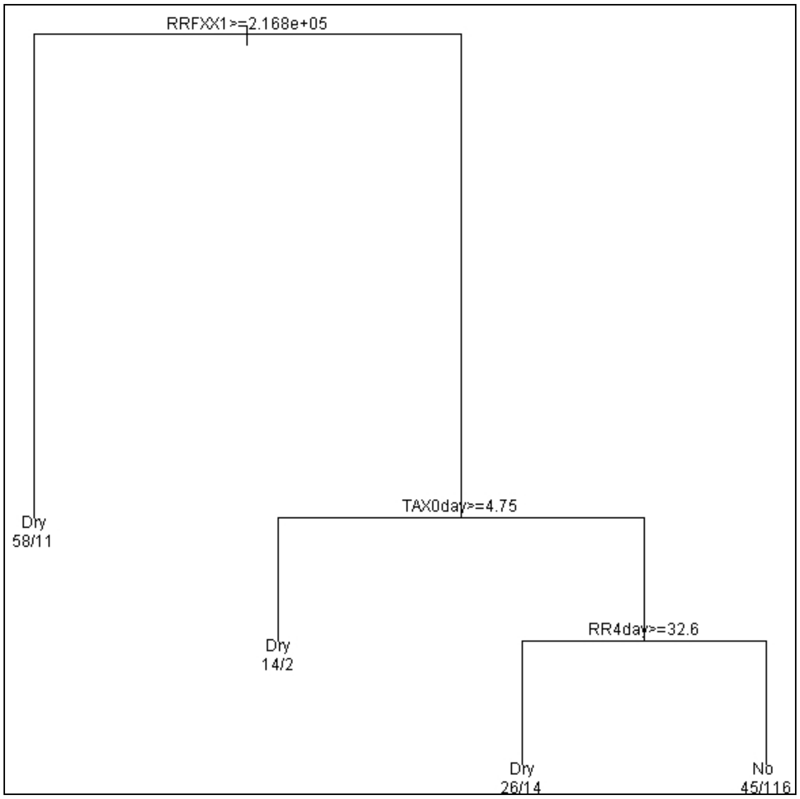
RUN 2



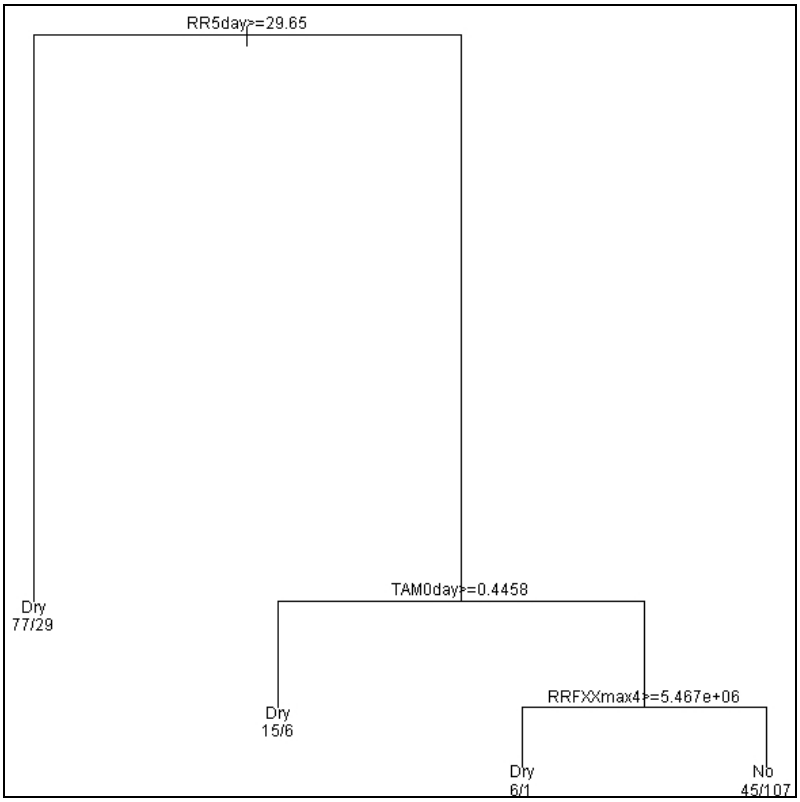
RUN 3


RR3day>=28.95
Dry
57/10
TAM0day>=-0.4
Dry
23/9
RRFXXmax1>=1.446e+06
Dry
9/1
FXX0day>=12.4
FFM0day< 9
No
20/73
Dry
14/4
No
20/46

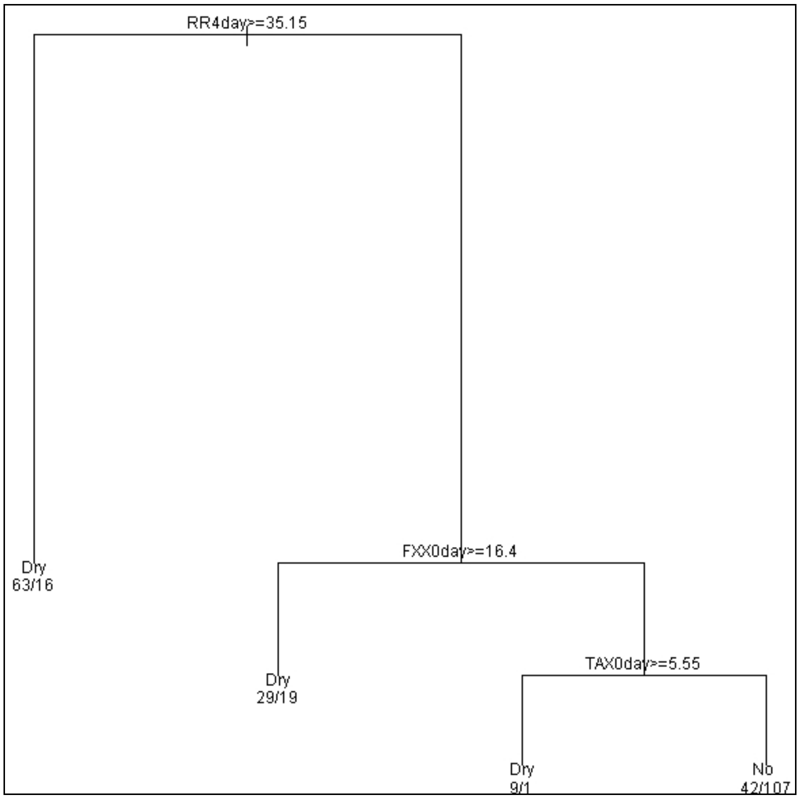
RUN 4


RRFXX1>=2.168e+05
Dry
58/11
TAX0day>=4.75
Dry
14/2
RR4day>=32.6
Dry
26/14
No
45/116

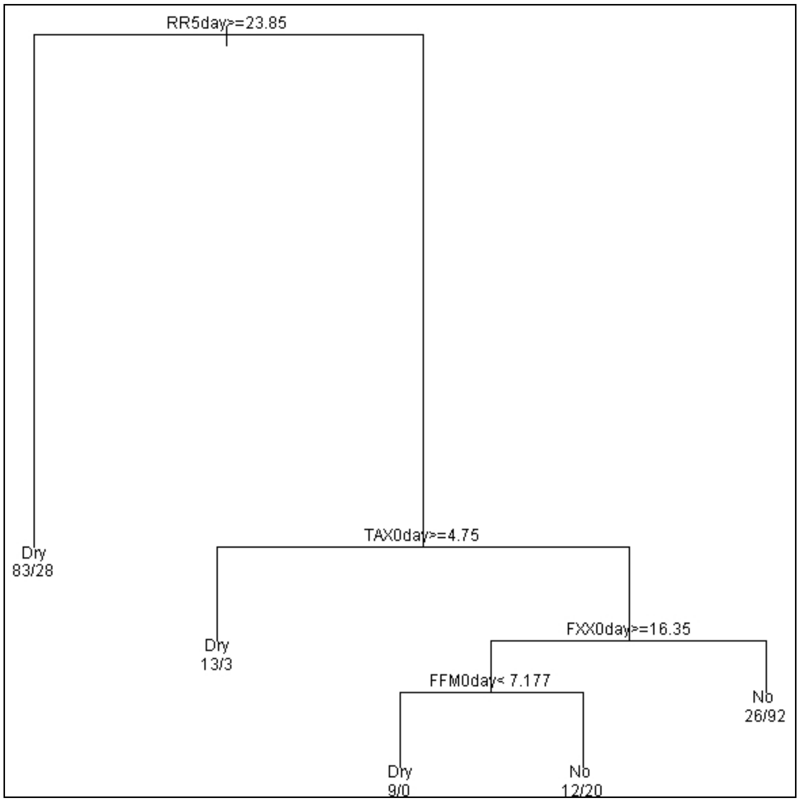
RUN 5


RR5day>=29.65
TAM0day>=0.4458
RRFXXmax4>=5.467e+06
Dry
77/29
Dry
15/6
Dry
6/1
No
45/107

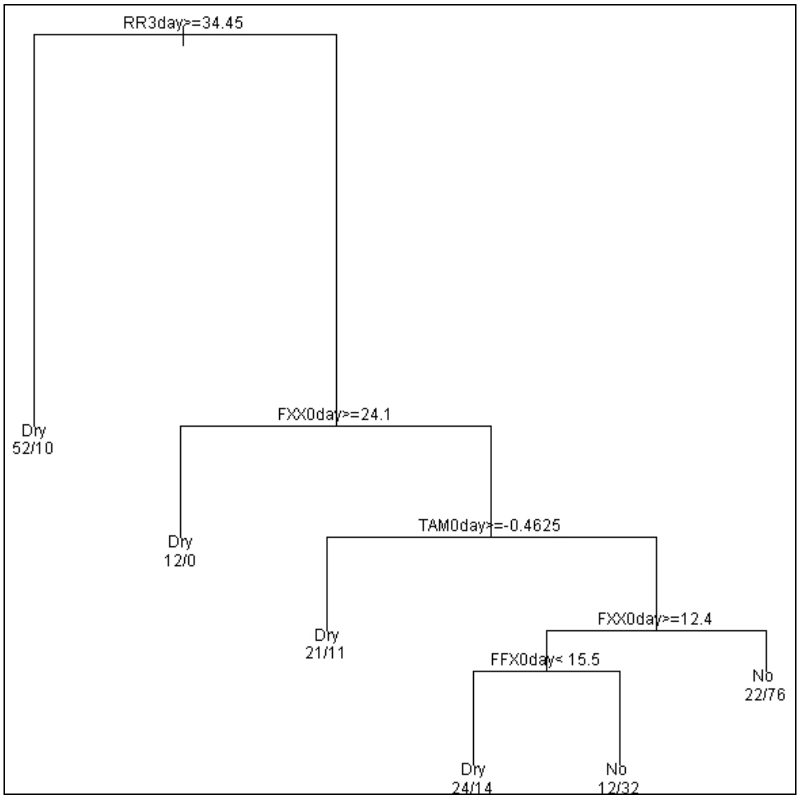
RUN 6


RR4day>=35.15
FXX0day>=16.4
TAX0day>=5.55
Dry
63/16
Dry
29/19
Dry
9/1
No
42/107

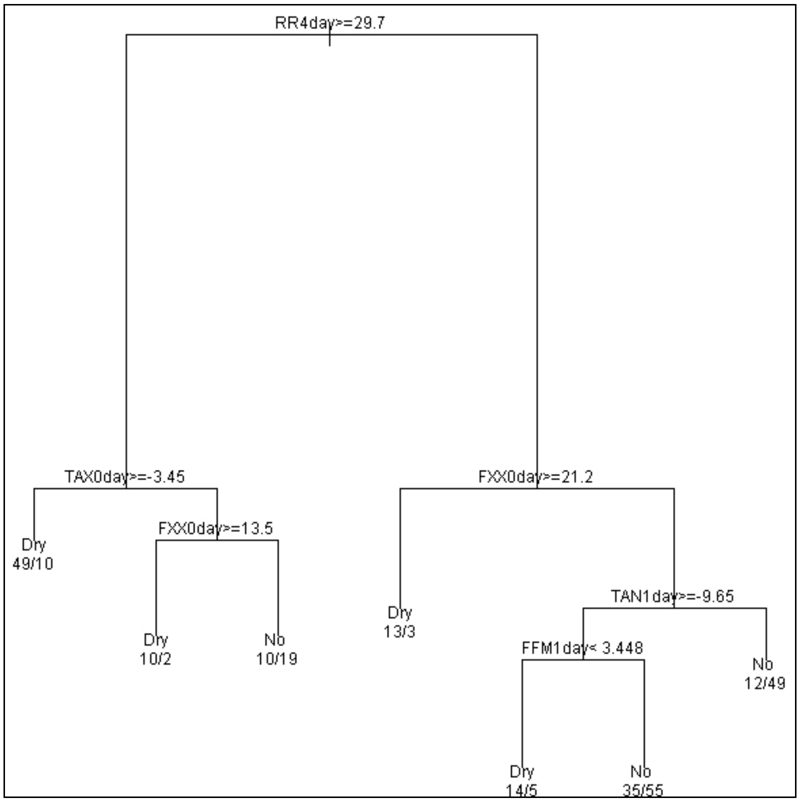
RUN 7


RR5day>=23.85
Dry
83/28
TAX0day>=4.75
Dry
13/3
FXX0day>=16.35
FFM0day< 7.177
No
26/92
Dry
9/0
No
12/20

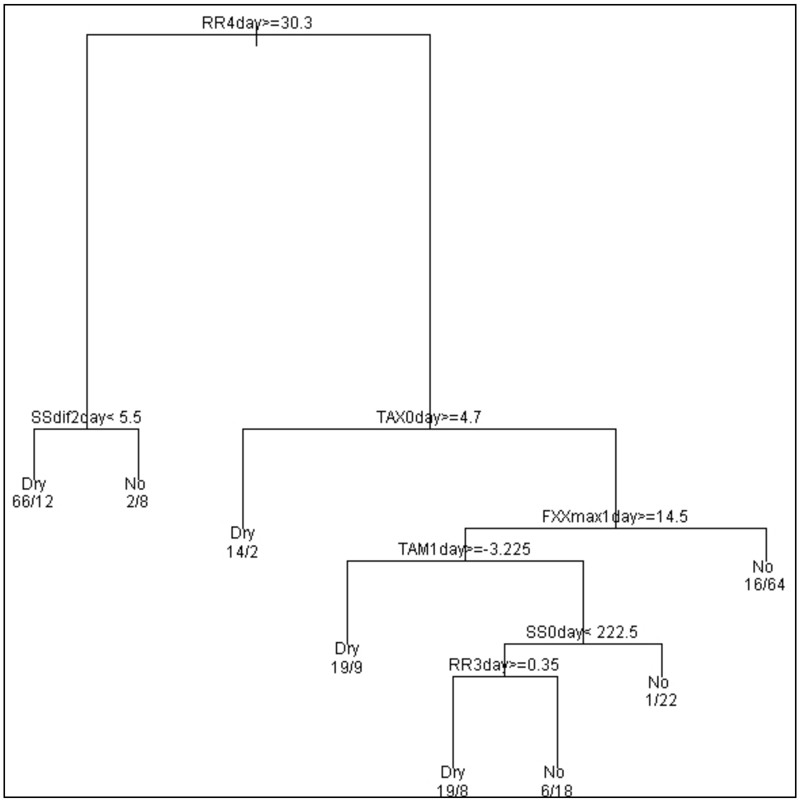
RUN 8


RR3day>=34.45
Dry
52/10
FXX0day>=24.1
Dry
12/0
TAM0day>=-0.4625
Dry
21/11
FXX0day>=12.4
FFX0day< 15.5
No
22/76
Dry
24/14
No
12/32

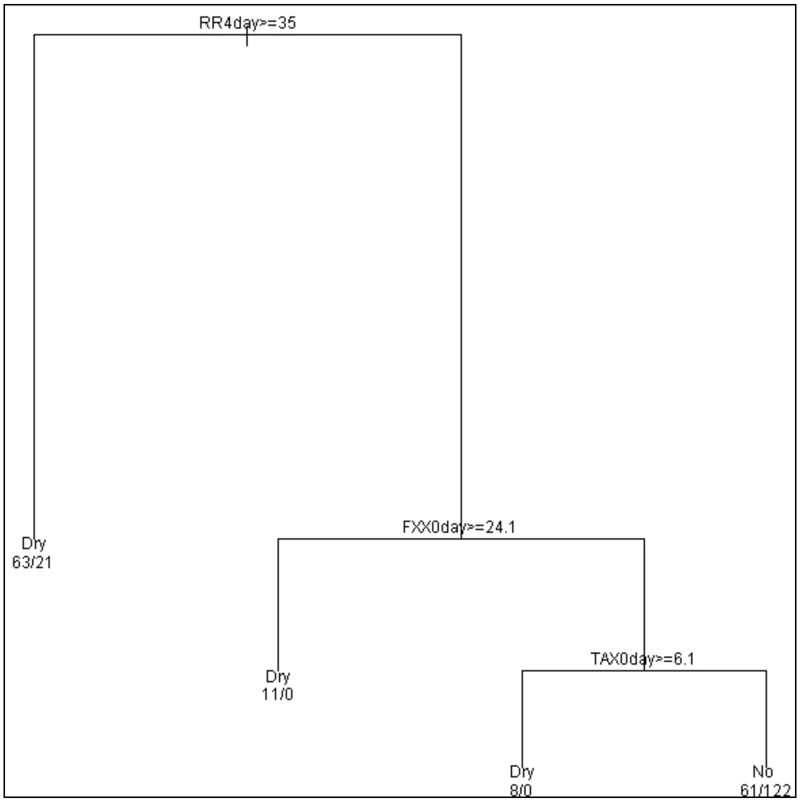
RUN 9


RR4day>=29.7
TAX0day>=-3.45
FXX0day>=21.2
Dry
49/10
FXX0day>=13.5
Dry
10/2
No
10/19
Dry
13/3
TAN1day>=-9.65
FFM1day< 3.448
No
12/49
Dry
14/5
No
35/55

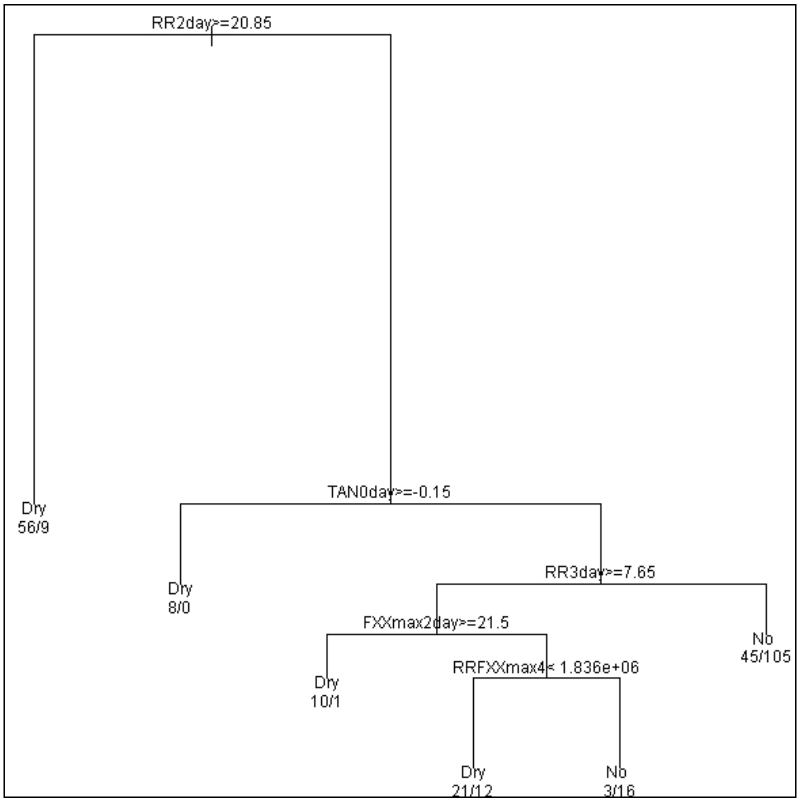
RUN10


RR4day>=30.3
SSdif2day< 5.5
TAX0day>=4.7
Dry
66/12
No
2/8
Dry
14/2
FXXmax1day>=14.5
TAM1day>=-3.225
No
16/64
Dry
19/9
SS0day< 222.5
RR3day>=0.35
No
1/22
Dry
19/8
No
6/18

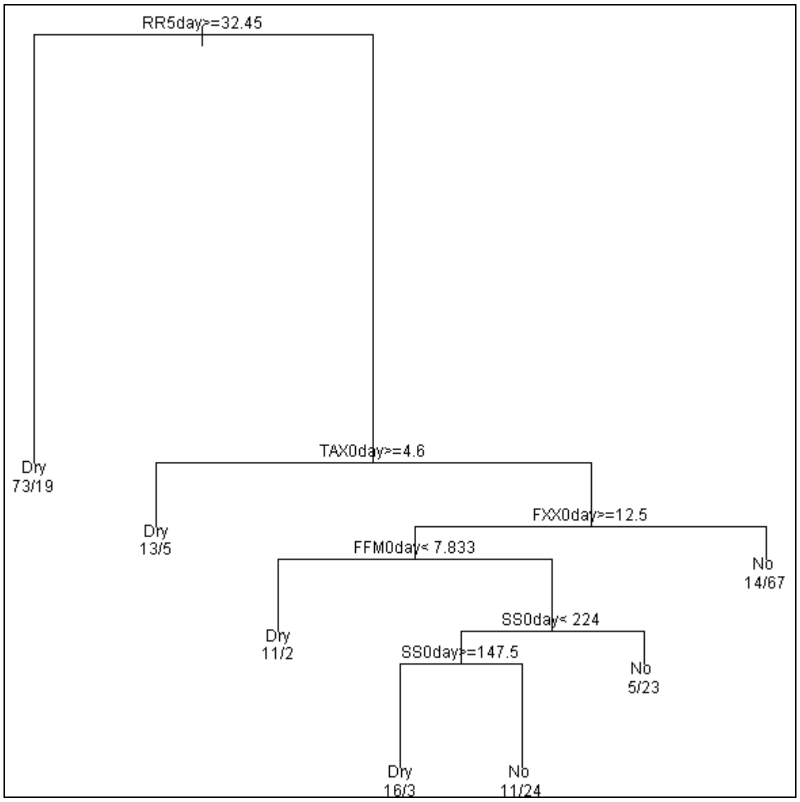
RUN 11


RR4day>=35
Dry
63/21
FXX0day>=24.1
Dry
11/0
TAX0day>=6.1
Dry
8/0
No
61/122

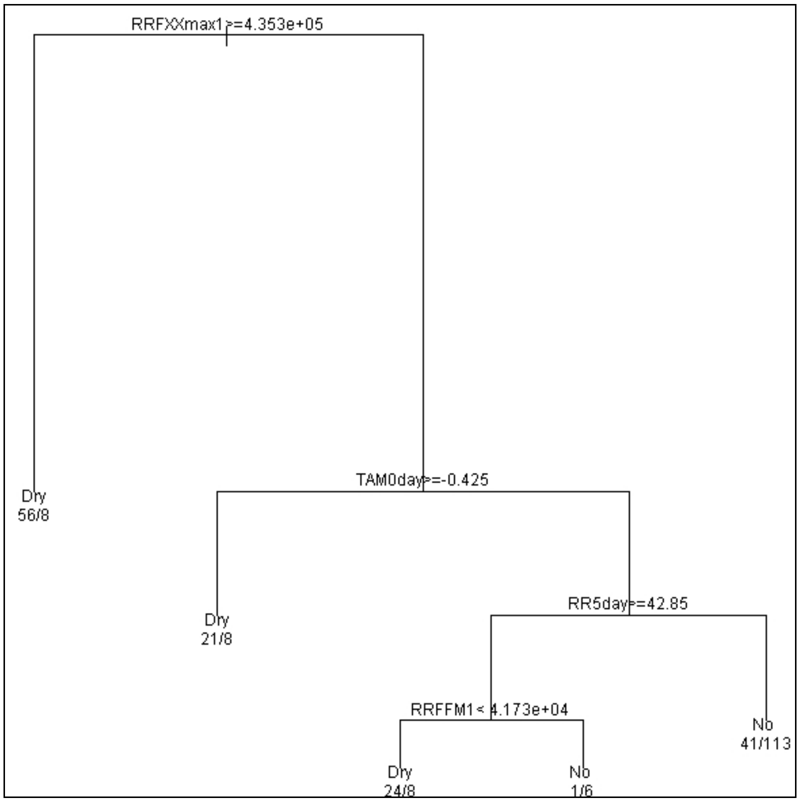
RUN 12


RR2day>=20.85
Dry
56/9
TAN0day>=-0.15
Dry
8/0
RR3day>=7.65
FXXmax2day>=21.5
No
45/105
Dry
10/1
RRFXXmax4< 1.836e+06
Dry
21/12
No
3/16

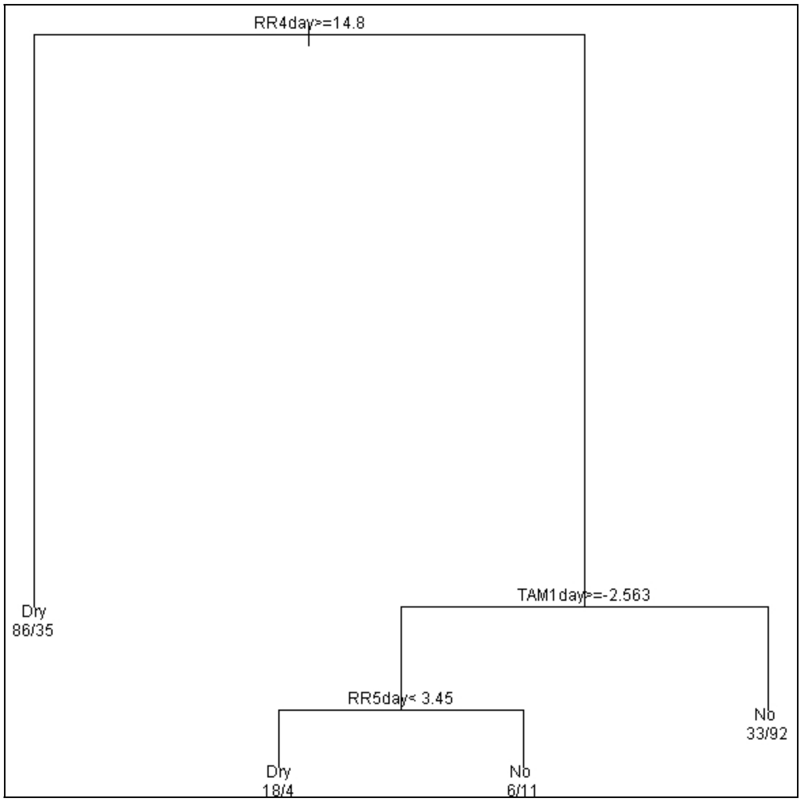
RUN 13


RR5day>=32.45
Dry
73/19
TAX0day>=4.6
Dry
13/5
FXX0day>=12.5
FFM0day< 7.833
No
14/67
Dry
11/2
SS0day< 224
SS0day>=147.5
No
5/23
Dry
16/3
No
11/24

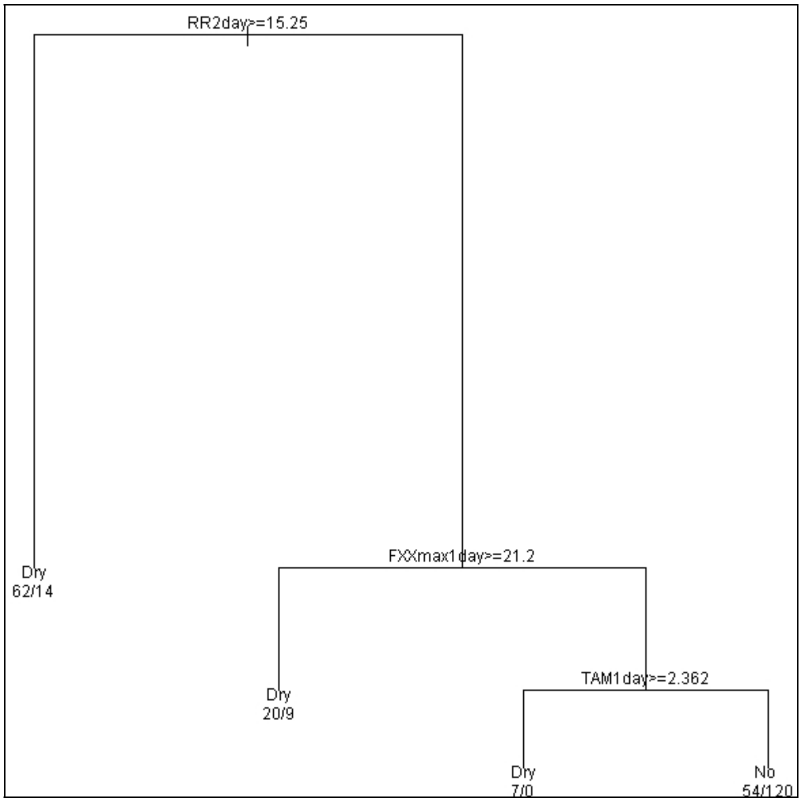
RUN 14


RRFXXmax1>=4.353e+05
Dry
56/8
TAM0day>=-0.425
Dry
21/8
RR5day>=42.85
RRFFM1< 4.173e+04
No
41/113
Dry
24/8
No
1/6

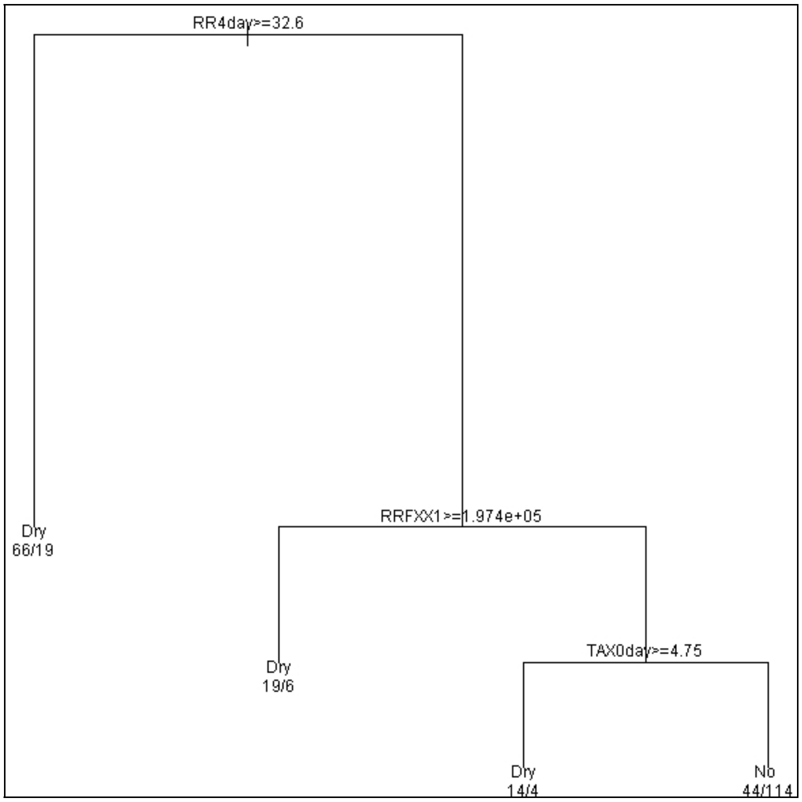
RUN 15


RR4day>=14.8
Dry
86/35
TAM1day>=-2.563
RR5day< 3.45
No
33/92
Dry
18/4
No
6/11

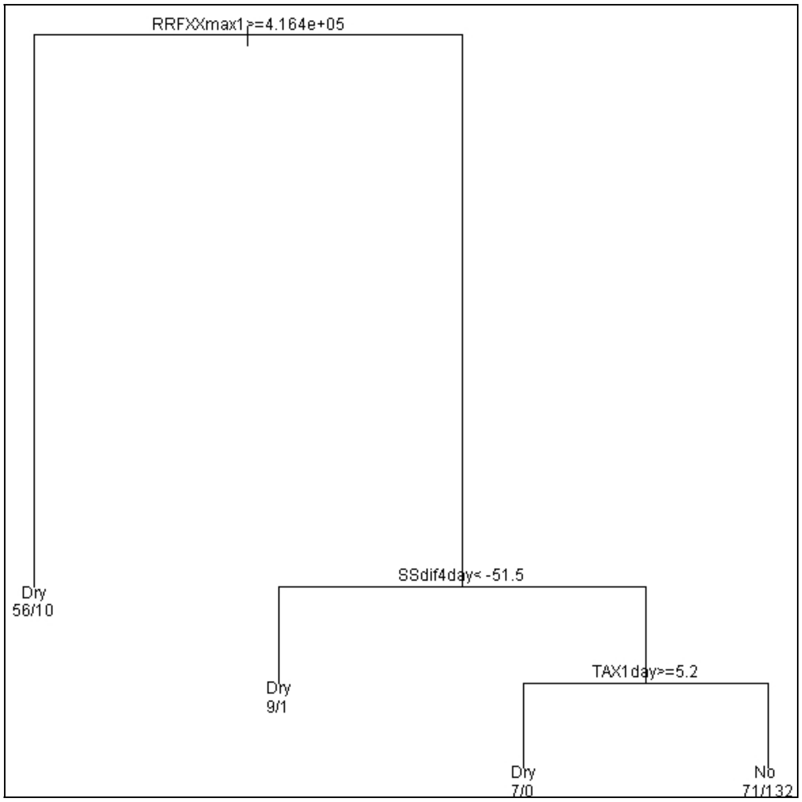
RUN 16


RR2day>=15.25
Dry
62/14
FXXmax1day>=21.2
Dry
20/9
TAM1day>=2.362
Dry
7/0
No
54/120

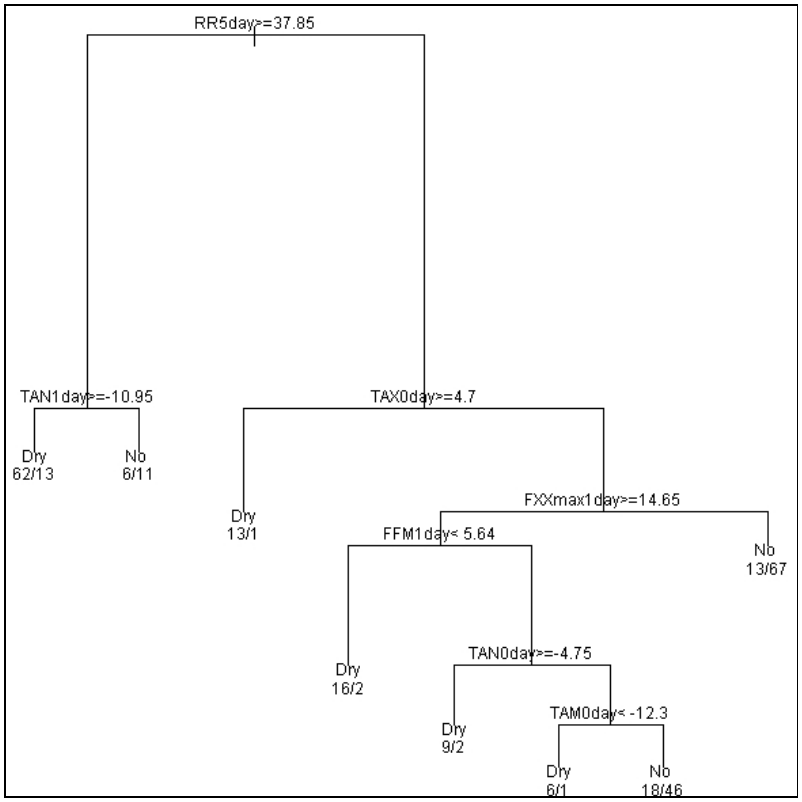
RUN 17


RR4day>=32.6
Dry
66/19
RRFXX1>=1.974e+05
Dry
19/6
TAX0day>=4.75
Dry
14/4
No
44/114

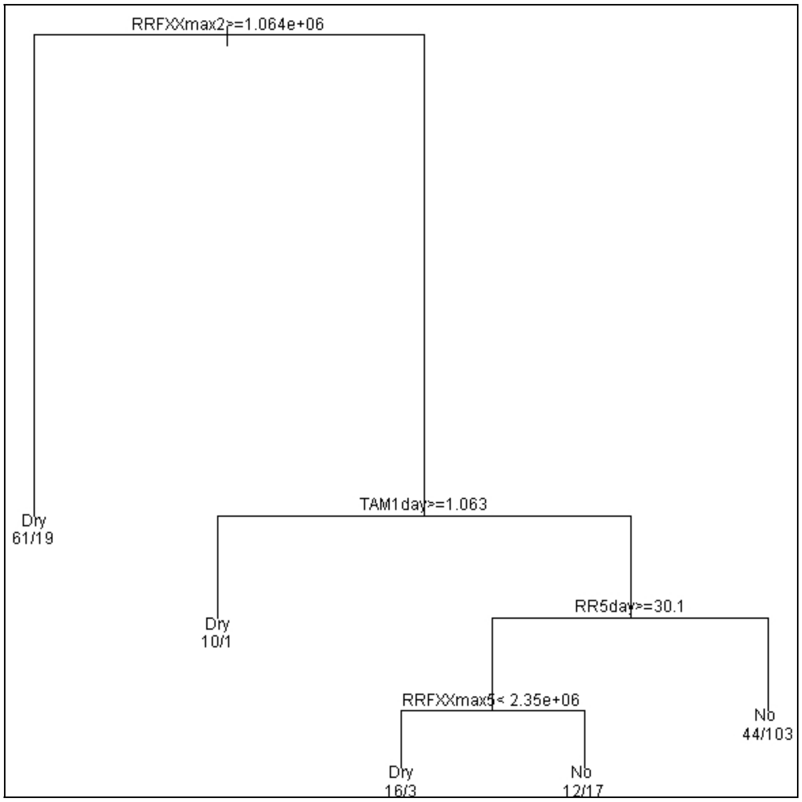
RUN 18


RRFXXmax1>=4.164e+05
Dry
56/10
SSdif4day< -51.5
Dry
9/1
TAX1day>=5.2
Dry
7/0
No
71/132

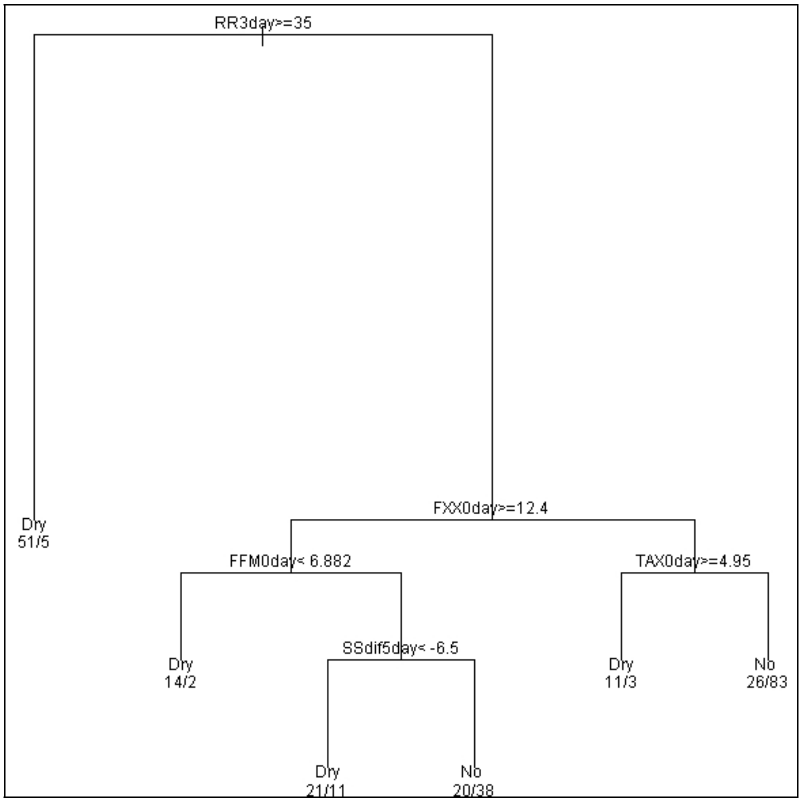
RUN 19


RR5day>=37.85
TAN1day>=-10.95
Dry
62/13
No
6/11
TAX0day>=4.7
Dry
13/1
FXXmax1day>=14.65
FFM1day< 5.64
No
13/67
Dry
16/2
TAN0day>=-4.75
Dry
9/2
TAM0day< -12.3
Dry
6/1
No
18/46

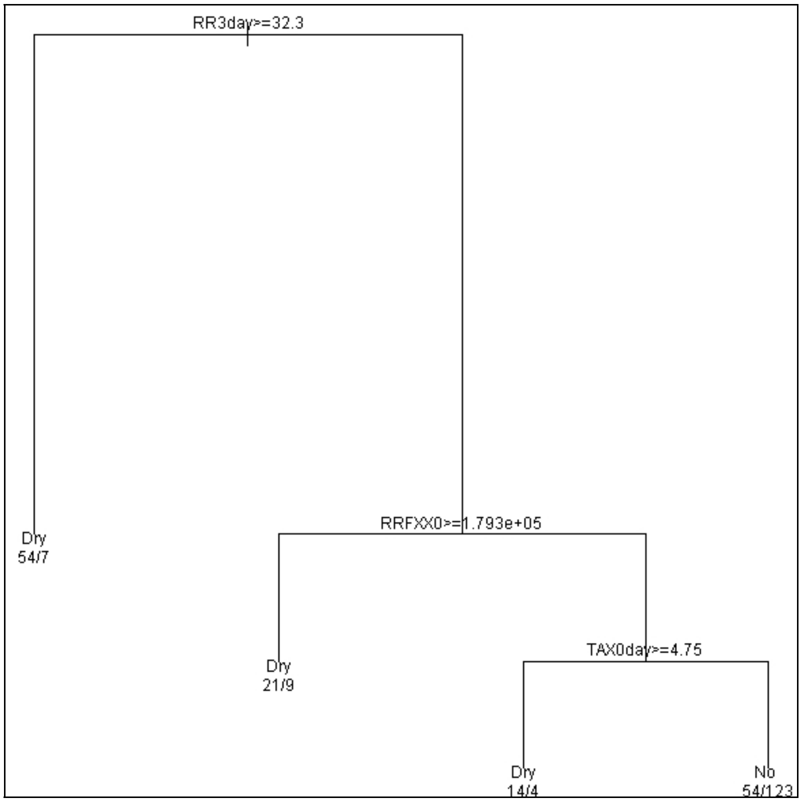
RUN 20


RRFXXmax2>=1.064e+06
Dry
61/19
TAM1day>=1.063
Dry
10/1
RR5day>=30.1
RRFXXmax5< 2.35e+06
No
44/103
Dry
16/3
No
12/17

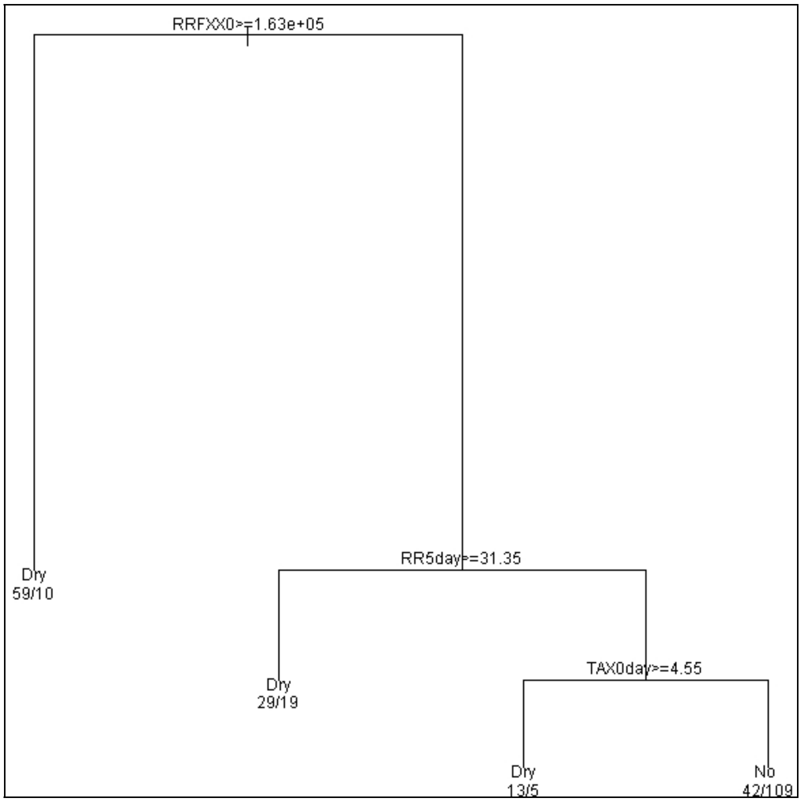
RUN 21


RR3day>=35
FXX0day>=12.4
Dry
51/5
FFM0day< 6.882
TAX0day>=4.95
SSdif5day< -6.5
Dry
14/2
Dry
11/3
No
26/83
Dry
21/11
No
20/38

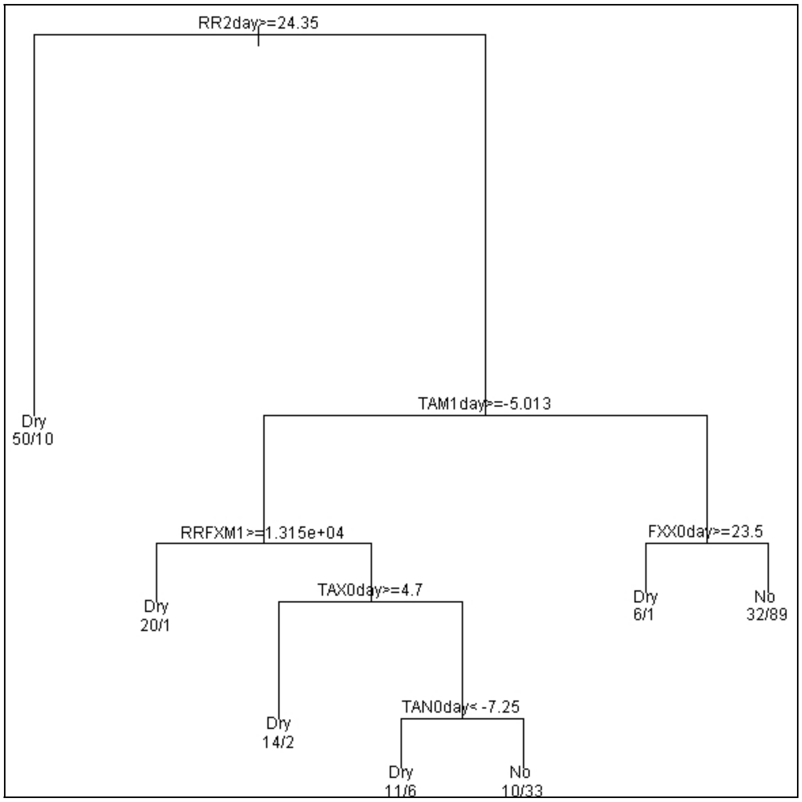
RUN 22


RR3day>=32.3
RRFXX0>=1.793e+05
Dry
54/7
Dry
21/9
TAX0day>=4.75
Dry
14/4
No
54/123

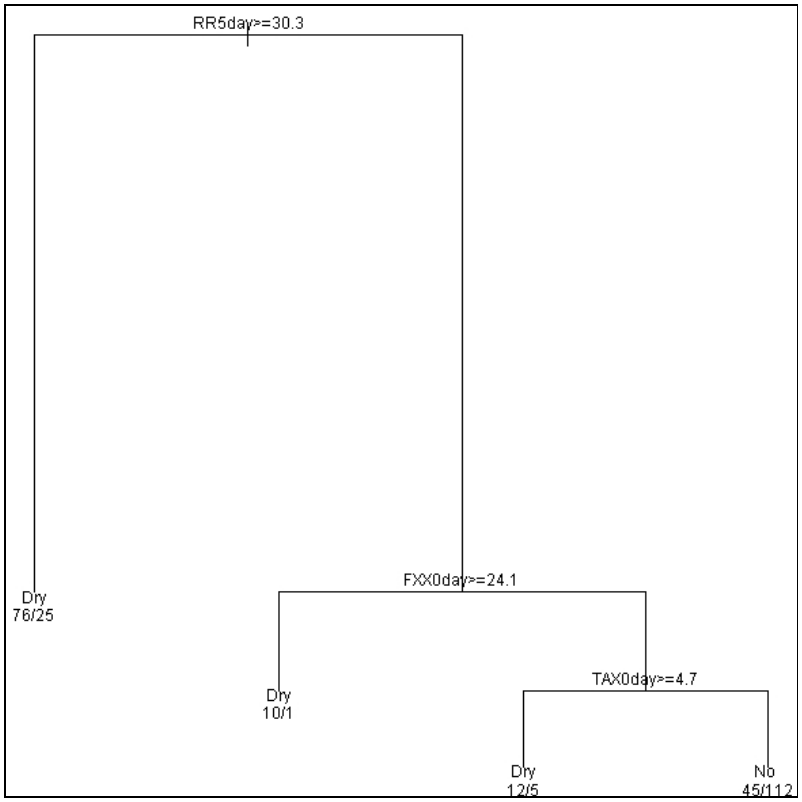
RUN 23


RRFXX0>=1.63e+05
Dry
59/10
RR5day>=31.35
Dry
29/19
TAX0day>=4.55
Dry
13/5
No
42/109

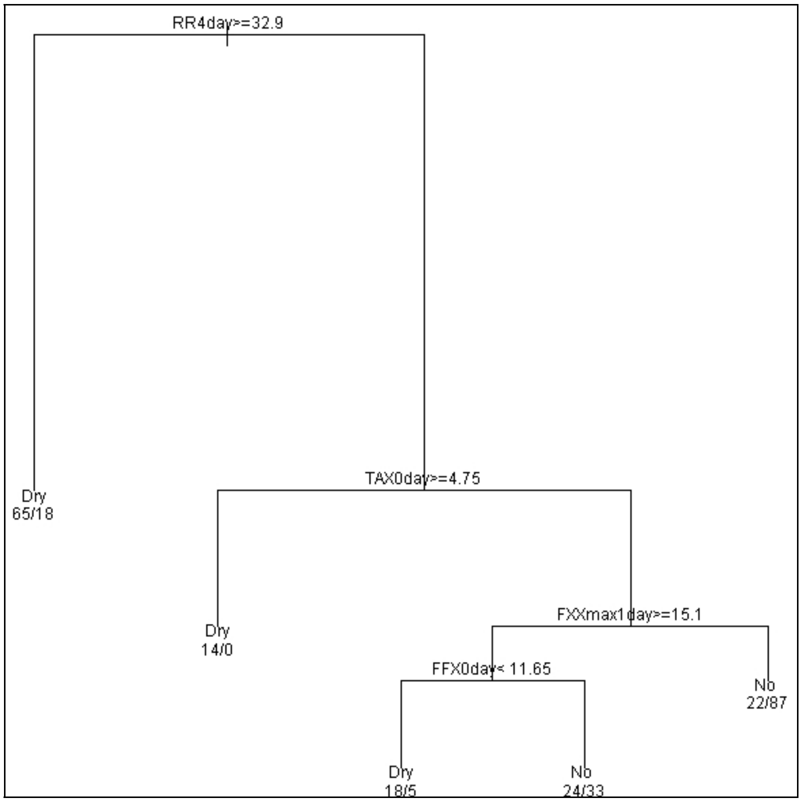
RUN 24


RR2day>=24.35
Dry
50/10
TAM1day>=-5.013
RRFXM1>=1.315e+04
Dry
20/1
TAX0day>=4.7
Dry
14/2
TAN0day< -7.25
Dry
11/6
No
10/33
FXX0day>=23.5
Dry
6/1
No
32/89

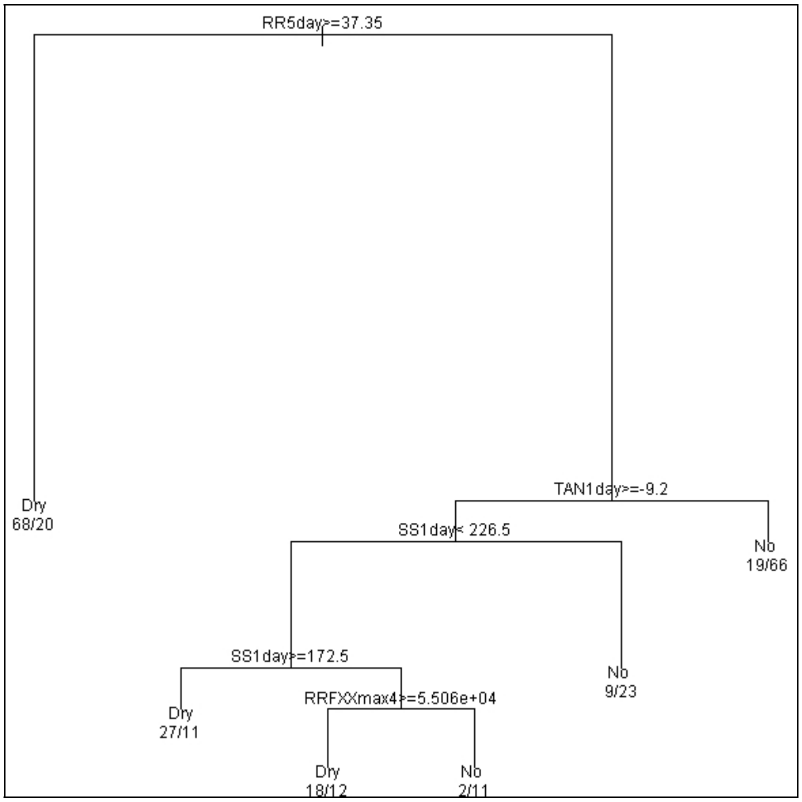
RUN 25


RR5day>=30.3
Dry
76/25
FXX0day>=24.1
Dry
10/1
TAX0day>=4.7
Dry
12/5
No
45/112

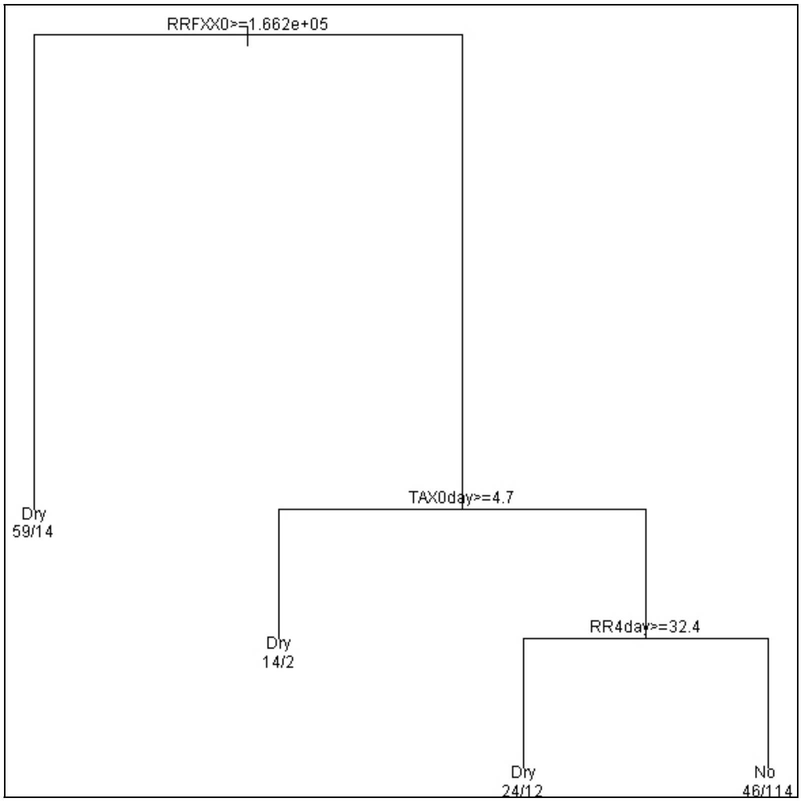
RUN 26


RR4day>=32.9
Dry
65/18
TAX0day>=4.75
Dry
14/0
FXXmax1day>=15.1
FFX0day< 11.65
No
22/87
Dry
18/5
No
24/33

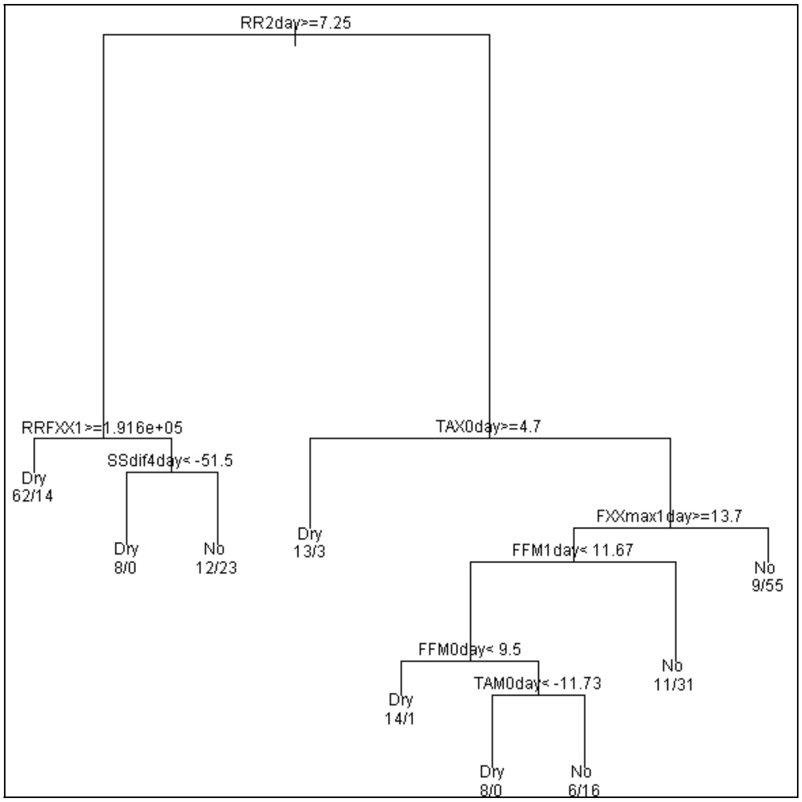
RUN 27


RR5day>=37.35
TAN1day>=-9.2
Dry
68/20
SS1day< 226.5
No
19/66
SS1day>=172.5
No
9/23
RRFXXmax4>=5.506e+04
Dry
27/11
Dry
18/12
No
2/11

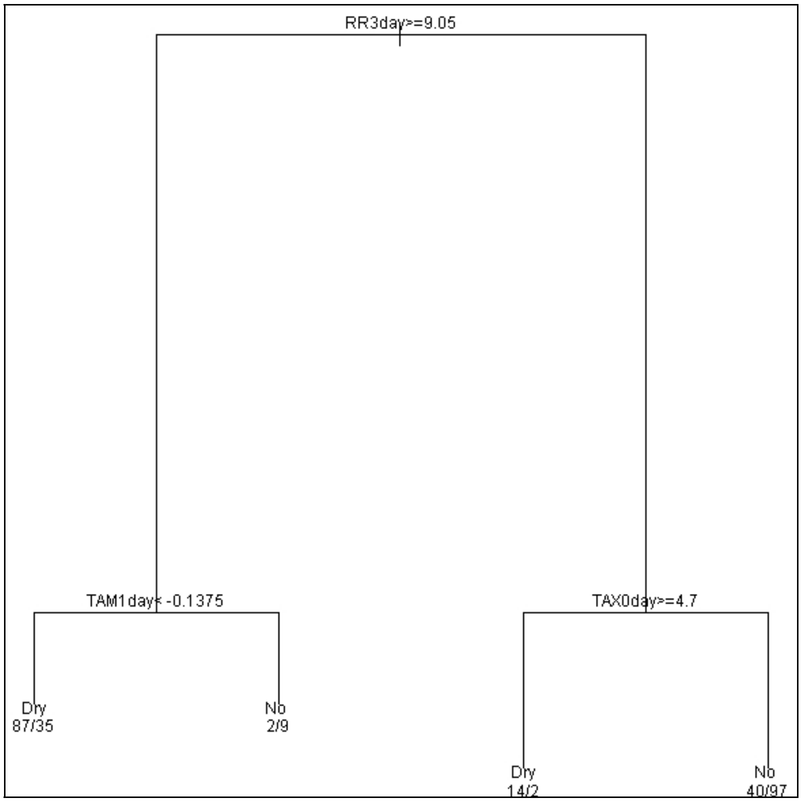
RUN 28


RRFXX0>=1.662e+05
TAX0day>=4.7
Dry
59/14
RR4day>=32.4
Dry
14/2
Dry
24/12
No
46/114

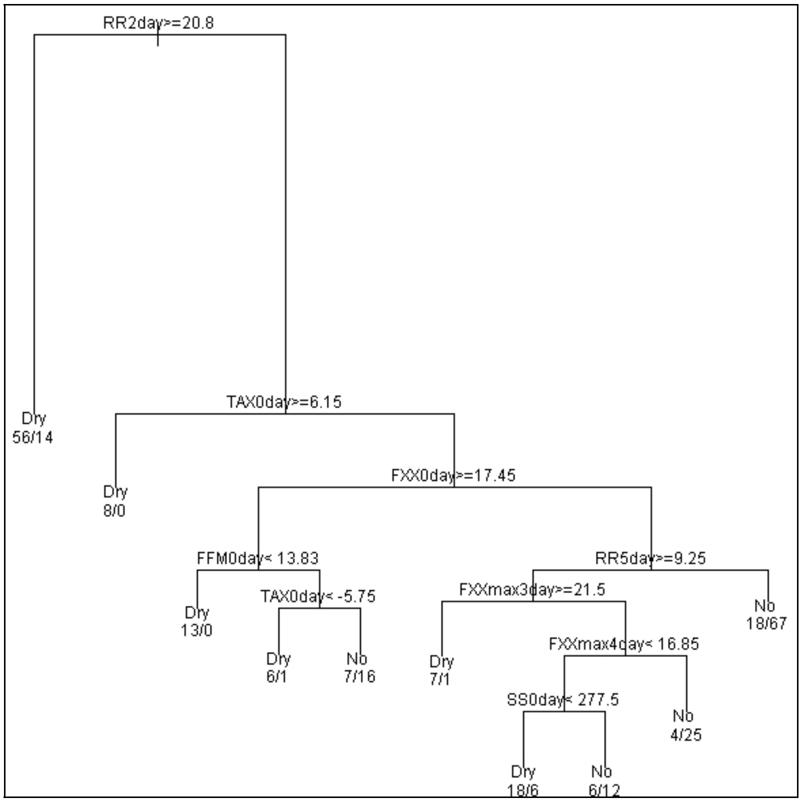
RUN 29


RR2day>=7.25
RRFXX1>=1.916e+05
Dry
62/14
SSdif4day< -51.5
Dry
8/0
No
12/23
TAX0day>=4.7
Dry
13/3
FXXmax1day>=13.7
FFM1day< 11.67
No
9/55
FFM0day< 9.5
Dry
14/1
TAM0day< -11.73
No
11/31
Dry
8/0
No
6/16

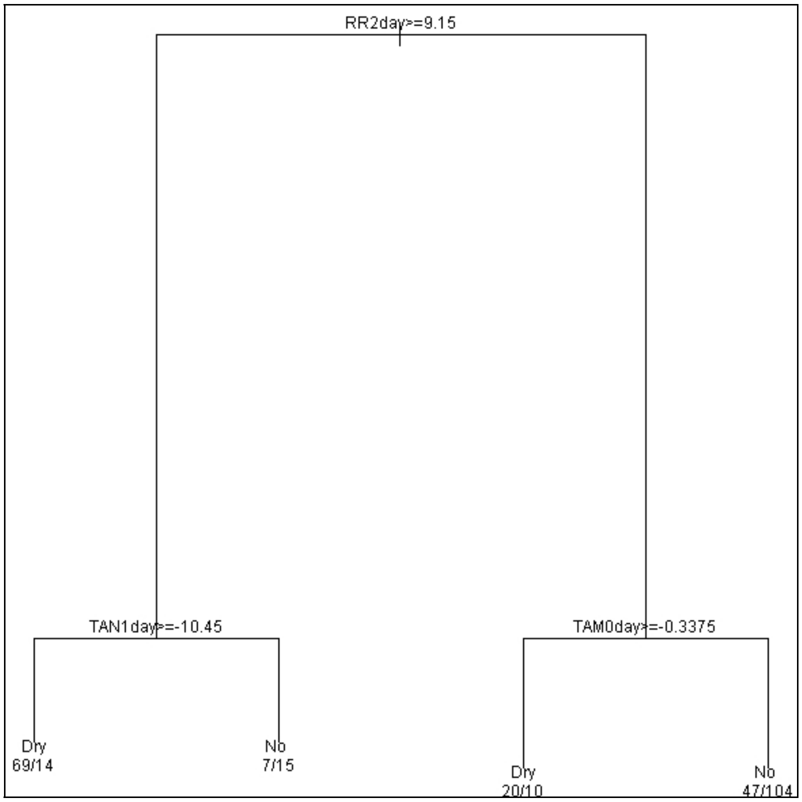
RUN 30


RR3day>=9.05
TAM1day< -0.1375
Dry
87/35
No
2/9
TAX0day>=4.7
Dry
14/2
No
40/97

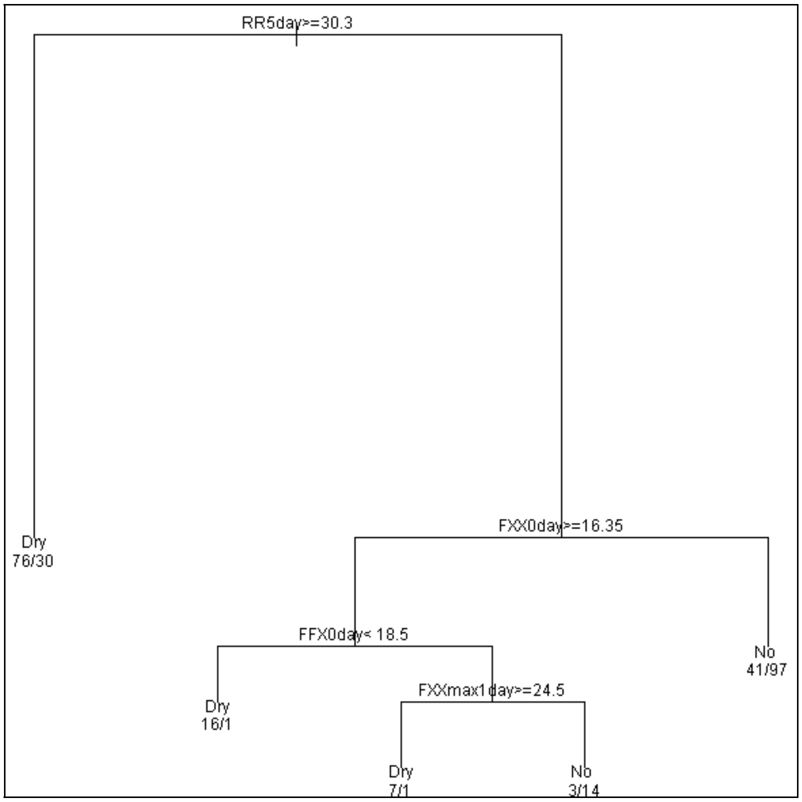
RUN 31


RR2day>=20.8
Dry
56/14
TAX0day>=6.15
Dry
8/0
FXX0day>=17.45
FFM0day< 13.83
Dry
13/0
TAX0day< -5.75
Dry
6/1
No
7/16
RR5day>=9.25
FXXmax3day>=21.5
Dry
7/1
FXXmax4day< 16.85
SS0day< 277.5
Dry
18/6
No
6/12
No
4/25
No
18/67

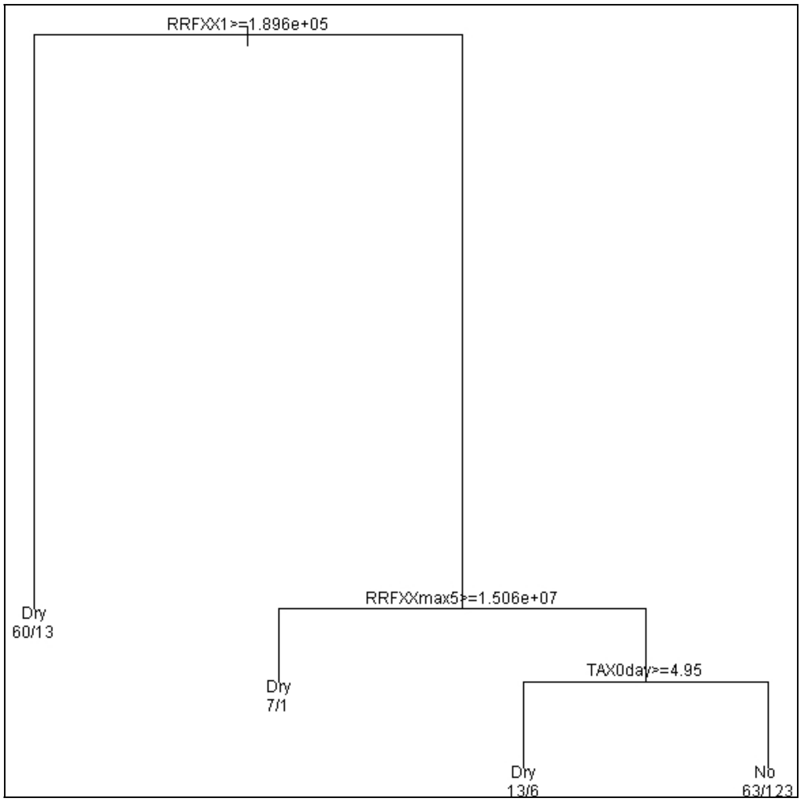
RUN 32


RR2day>=9.15
TAN1day>=-10.45
Dry
69/14
No
7/15
TAM0day>=-0.3375
Dry
20/10
No
47/104

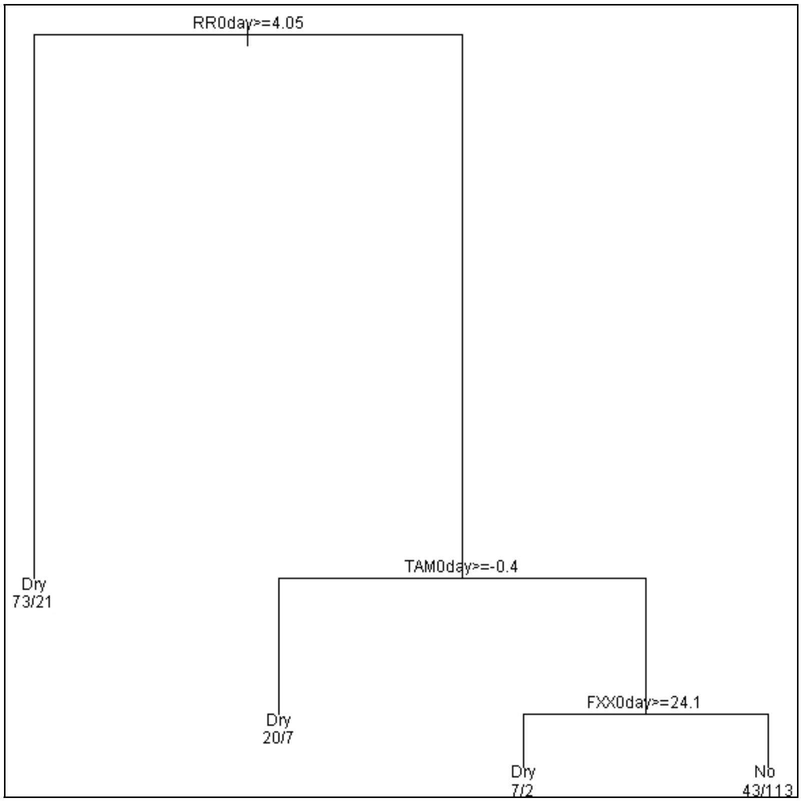
RUN 33


RR5day>=30.3
Dry
76/30
FXX0day>=16.35
FFX0day< 18.5
No
41/97
FXXmax1day>=24.5
Dry
16/1
Dry
7/1
No
3/14

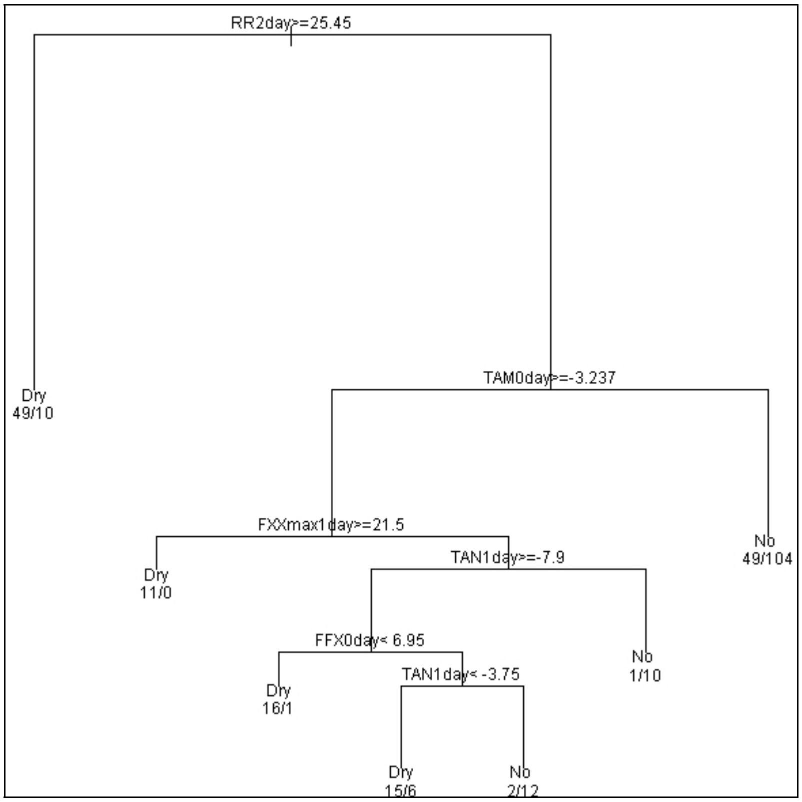
RUN 34


RRFXX1>=1.896e+05
Dry
60/13
RRFXXmax5>=1.506e+07
TAX0day>=4.95
Dry
7/1
Dry
13/6
No
63/123

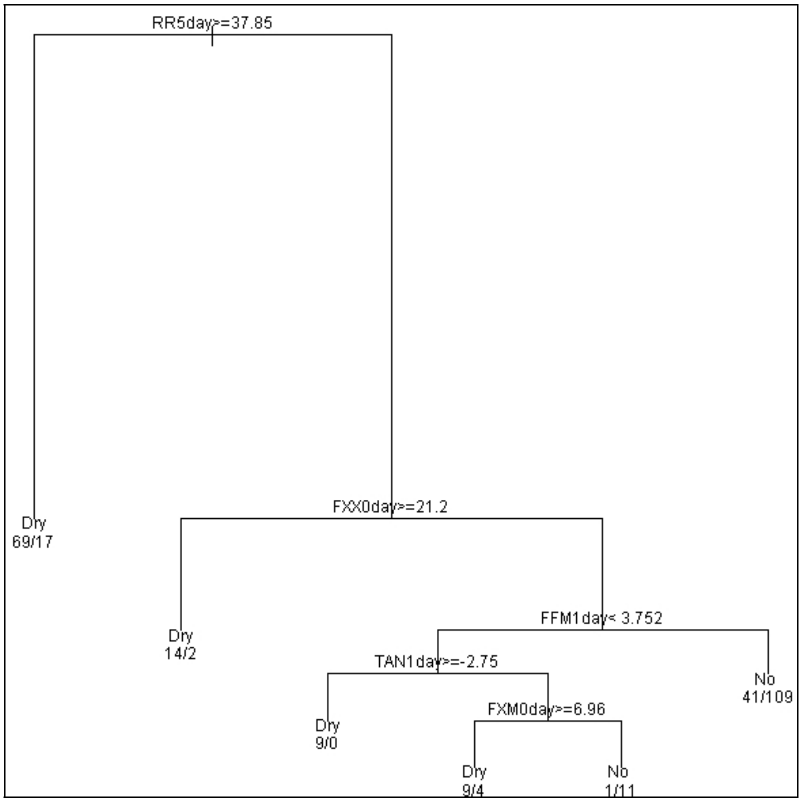
RUN 35


RR0day>=4.05
Dry
73/21
TAM0day>=-0.4
Dry
20/7
FXX0day>=24.1
Dry
7/2
No
43/113

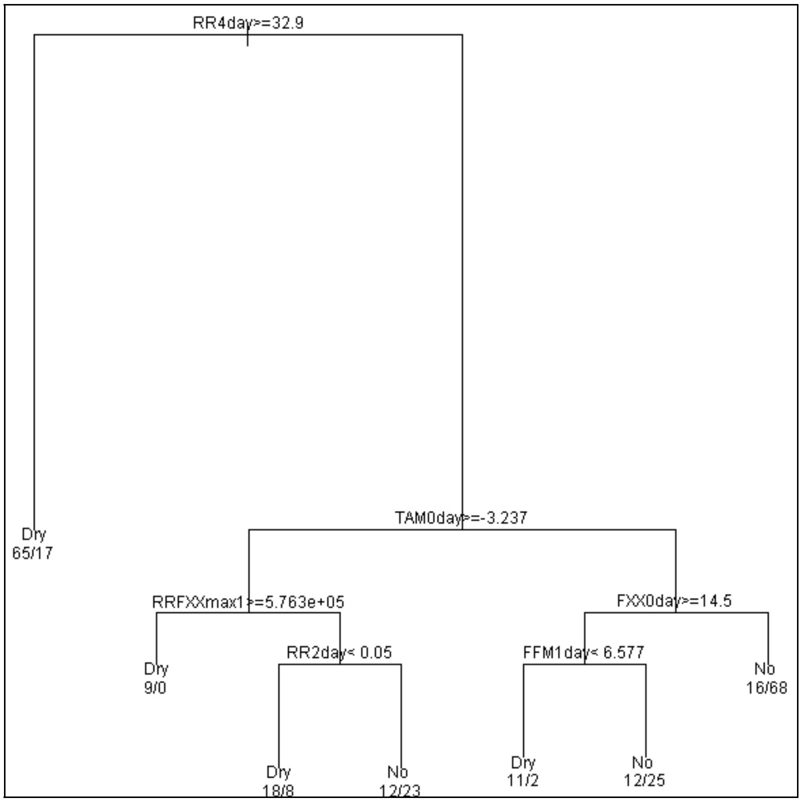
RUN 36


RR2day>=25.45
Dry
49/10
TAM0day>=-3.237
FXXmax1day>=21.5
Dry
11/0
TAN1day>=-7.9
No
49/104
FFX0day< 6.95
Dry
16/1
TAN1day< -3.75
No
1/10
Dry
15/6
No
2/12

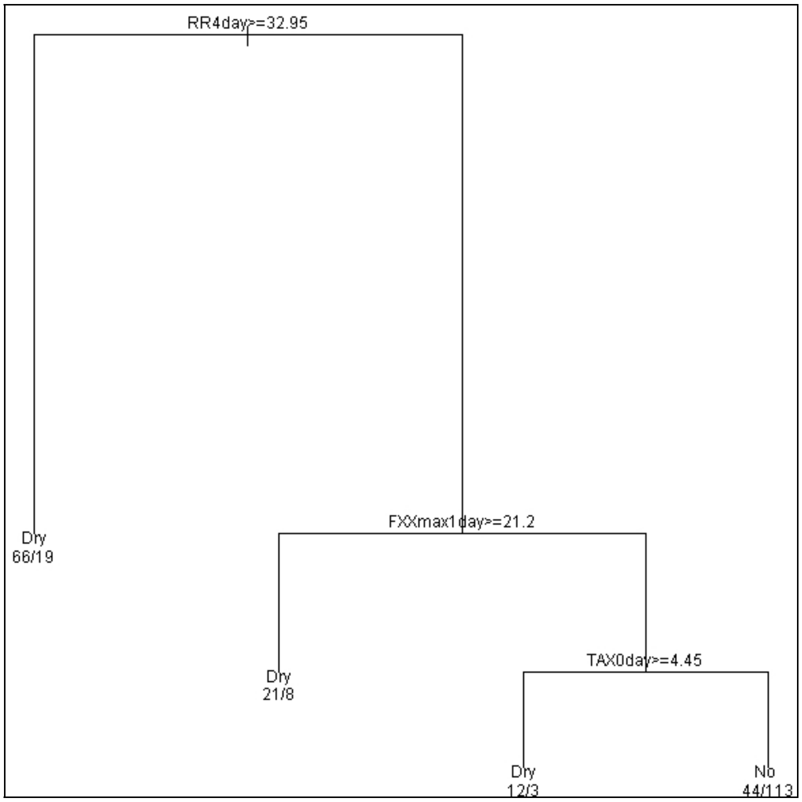
RUN 37


RR5day>=37.85
FXX0day>=21.2
Dry
69/17
Dry
14/2
FFM1day< 3.752
TAN1day>=-2.75
No
41/109
Dry
9/0
FXM0day>=6.96
Dry
9/4
No
1/11

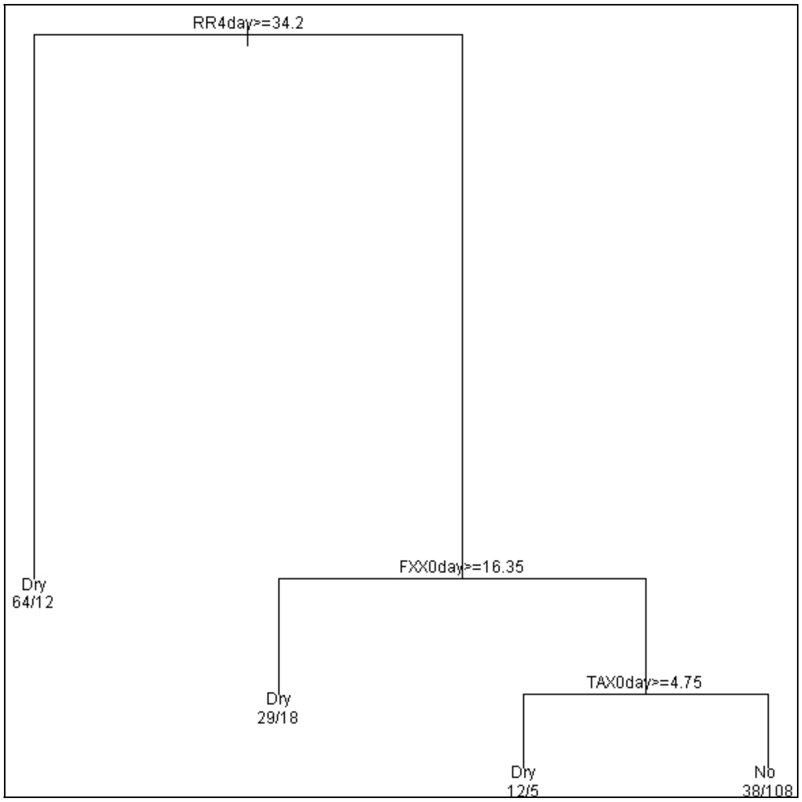
RUN 38


RR4day>=32.9
TAM0day>=-3.237
Dry
65/17
RRFXXmax1>=5.763e+05
FXX0day>=14.5
RR2day< 0.05
FFM1day< 6.577
Dry
9/0
No
16/68
Dry
18/8
No
12/23
Dry
11/2
No
12/25

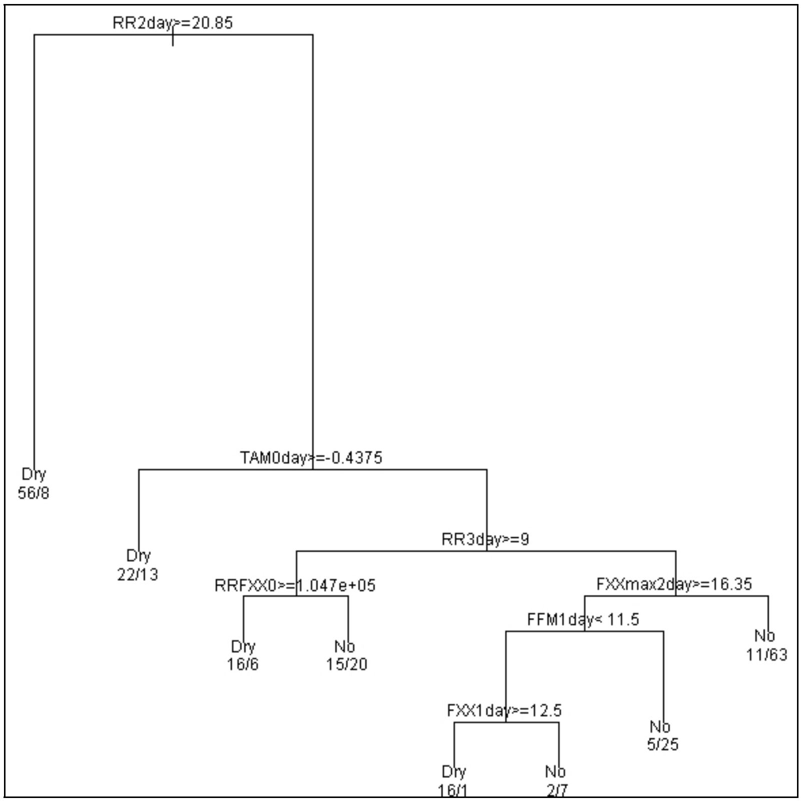
RUN 39


RR4day>=32.95
FXXmax1day>=21.2
TAX0day>=4.45
Dry
66/19
Dry
21/8
Dry
12/3
No
44/113

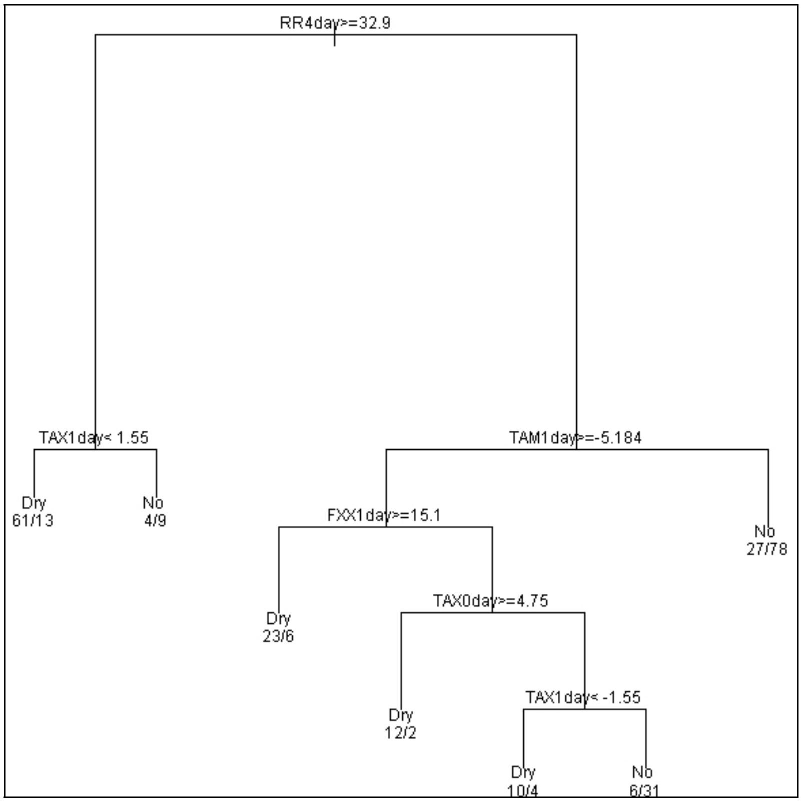
RUN 40


RR4day>=34.2
FXX0day>=16.35
TAX0day>=4.75
Dry
64/12
Dry
29/18
Dry
12/5
No
38/108

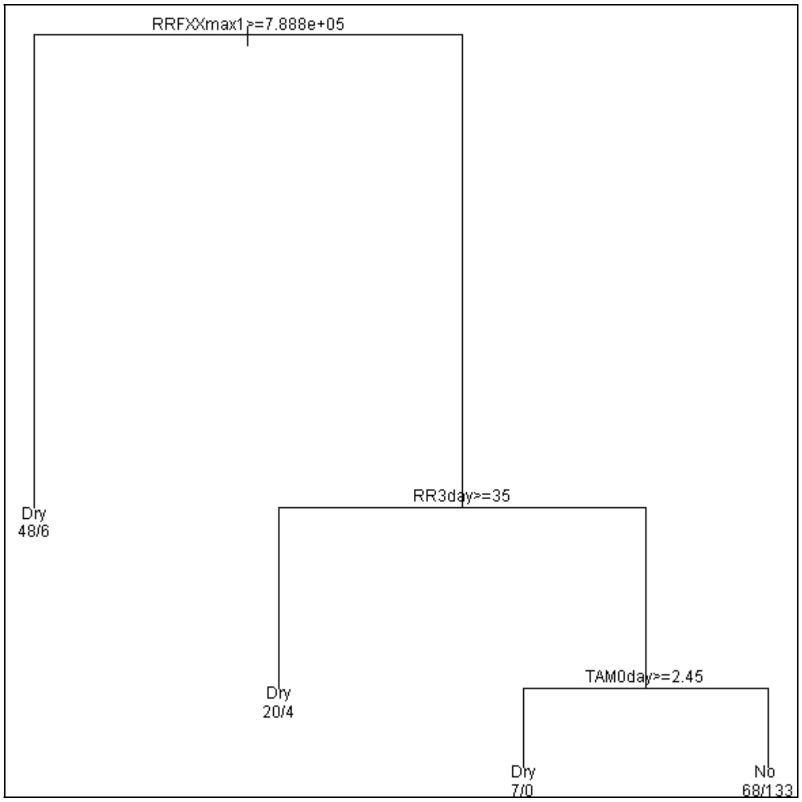
RUN 41


RR2day>=20.85
Dry
56/8
TAM0day>=-0.4375
Dry
22/13
RR3day>=9
RRFXX0>=1.047e+05
Dry
16/6
No
15/20
FXXmax2day>=16.35
FFM1day< 11.5
No
11/63
FXX1day>=12.5
No
5/25
Dry
16/1
No
2/7

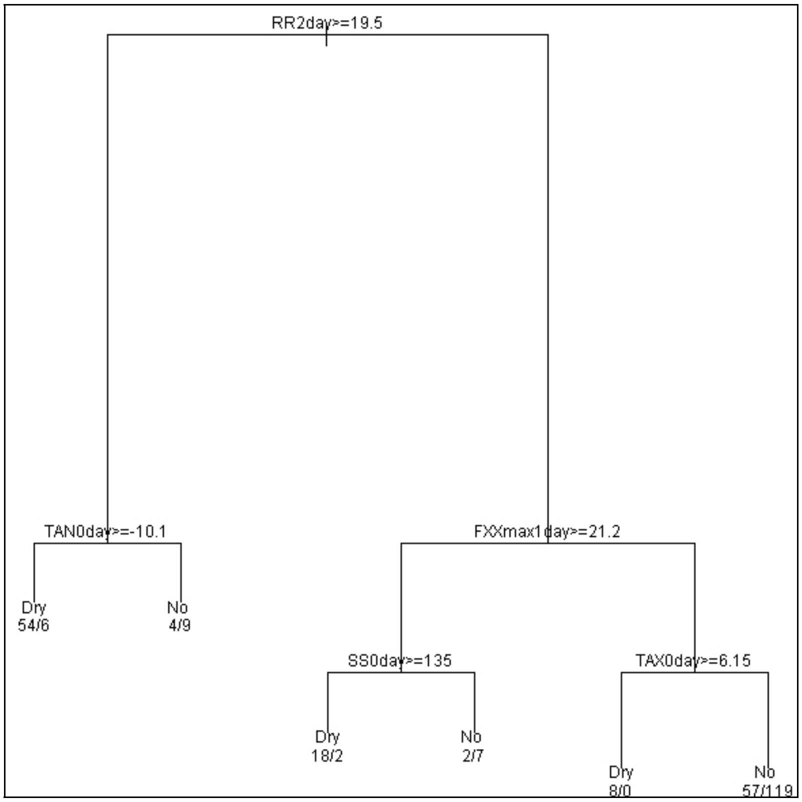
RUN 42


RR4day>=32.9
TAX1day< 1.55
Dry
61/13
No
4/9
TAM1day>=-5.184
FXX1day>=15.1
No
27/78
Dry
23/6
TAX0day>=4.75
Dry
12/2
TAX1day< -1.55
Dry
10/4
No
6/31

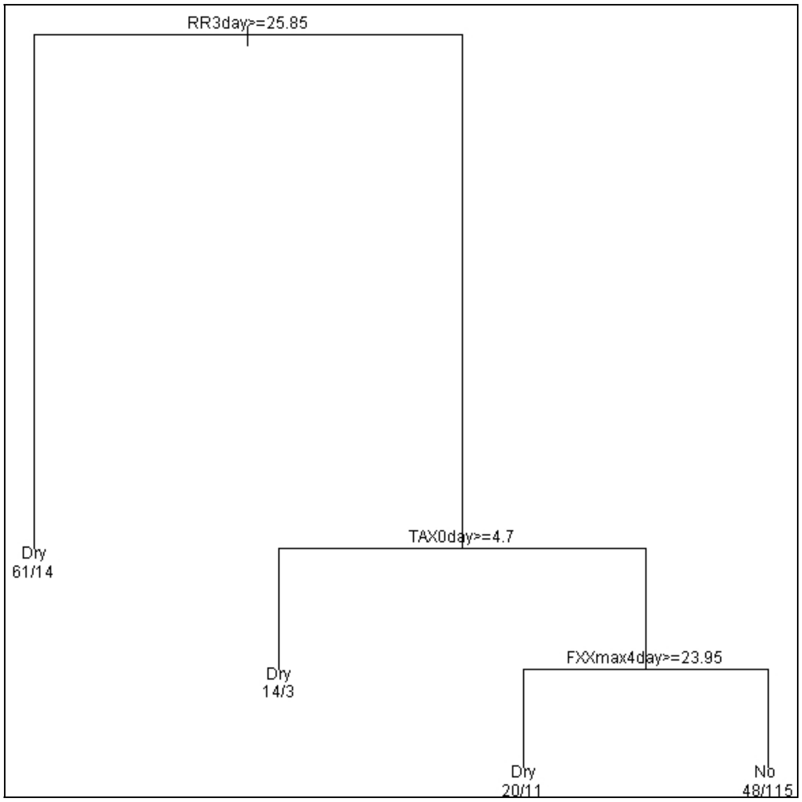
RUN 43


RRFXXmax1>=7.888e+05
Dry
48/6
RR3day>=35
Dry
20/4
TAM0day>=2.45
Dry
7/0
No
68/133

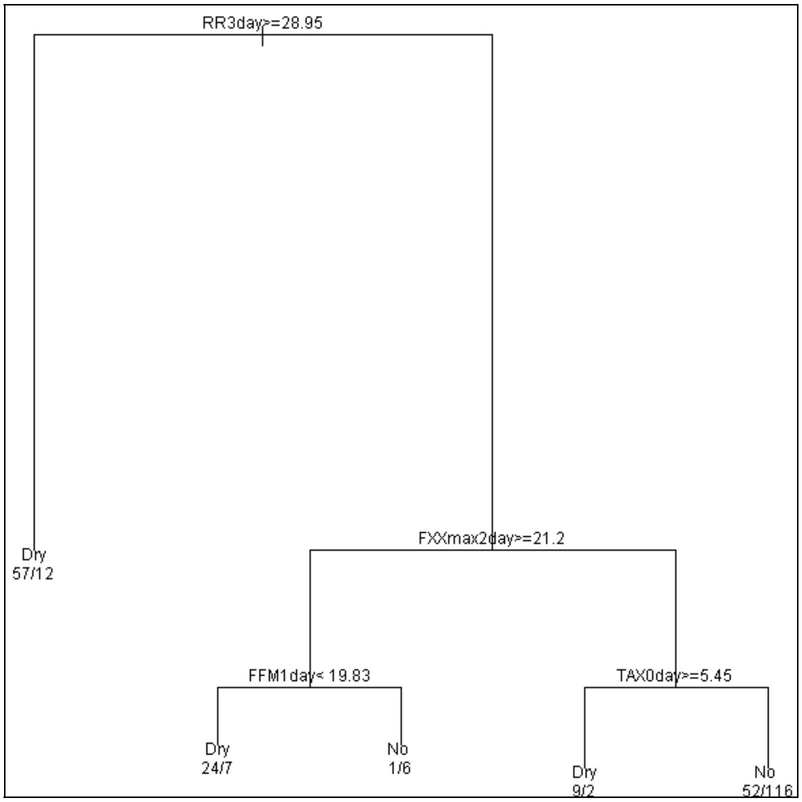
RUN 44


RR2day>=19.5
TAN0day>=-10.1
Dry
54/6
No
4/9
FXXmax1day>=21.2
SS0day>=135
Dry
18/2
No
2/7
TAX0day>=6.15
Dry
8/0
No
57/119

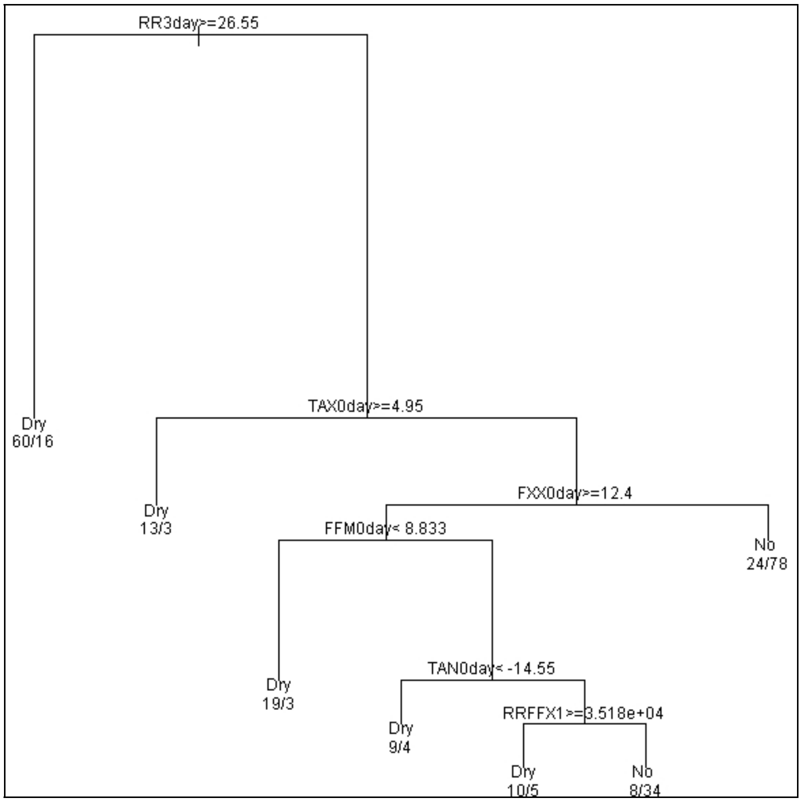
RUN 45


RR3day>=25.85
Dry
61/14
TAX0day>=4.7
Dry
14/3
FXXmax4day>=23.95
Dry
20/11
No
48/115

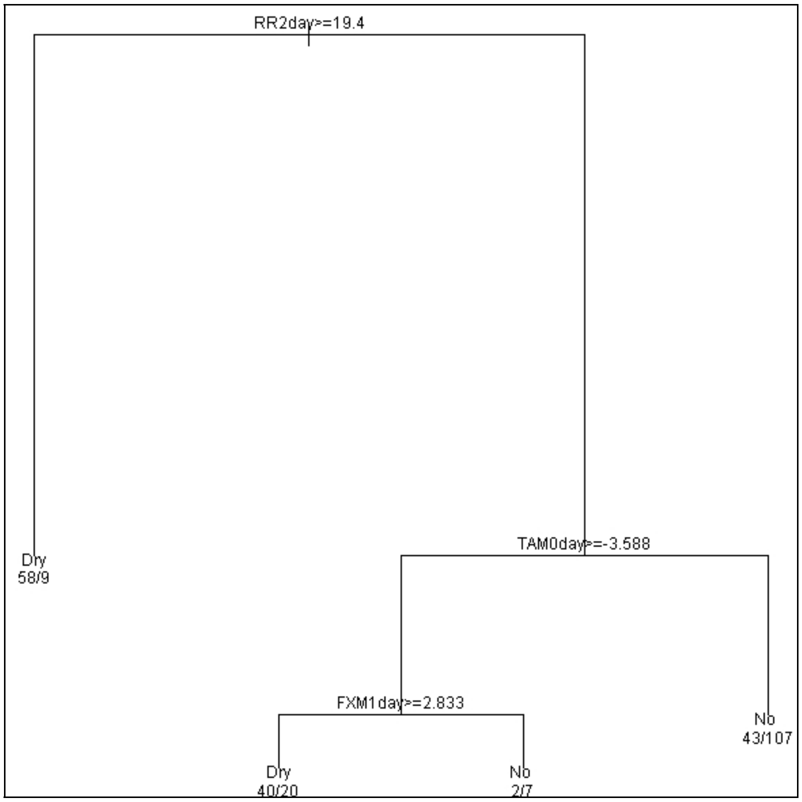
RUN 46


RR3day>=28.95
Dry
57/12
FXXmax2day>=21.2
FFM1day< 19.83
Dry
24/7
No
1/6
TAX0day>=5.45
Dry
9/2
No
52/116

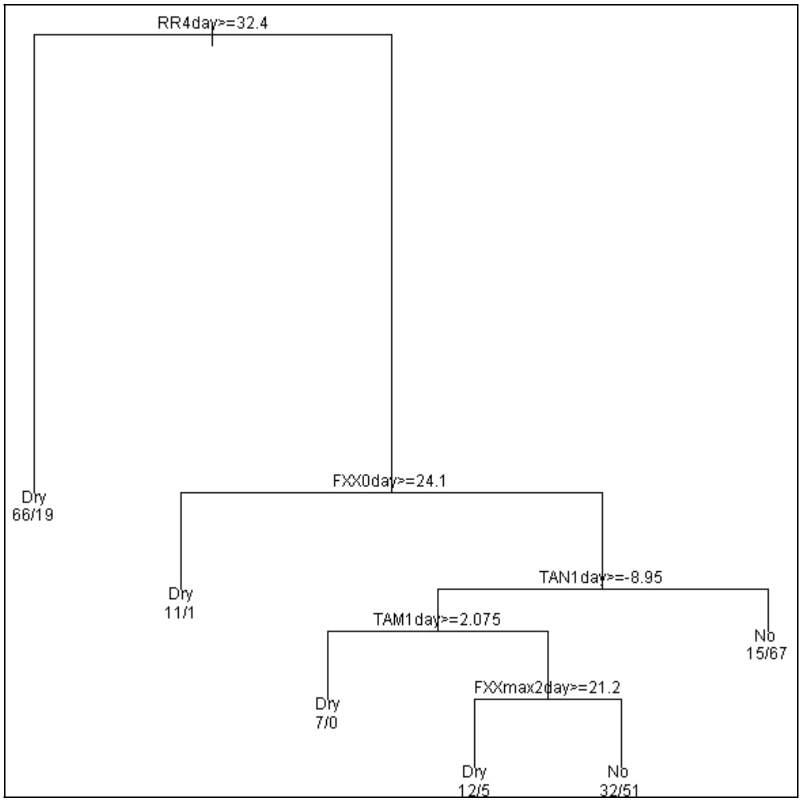
RUN 47


RR3day>=26.55
Dry
60/16
TAX0day>=4.95
Dry
13/3
FXX0day>=12.4
No
24/78
FFM0day< 8.833
Dry
19/3
TAN0day< -14.55
Dry
9/4
RRFFX1>=3.518e+04
Dry
10/5
No
8/34

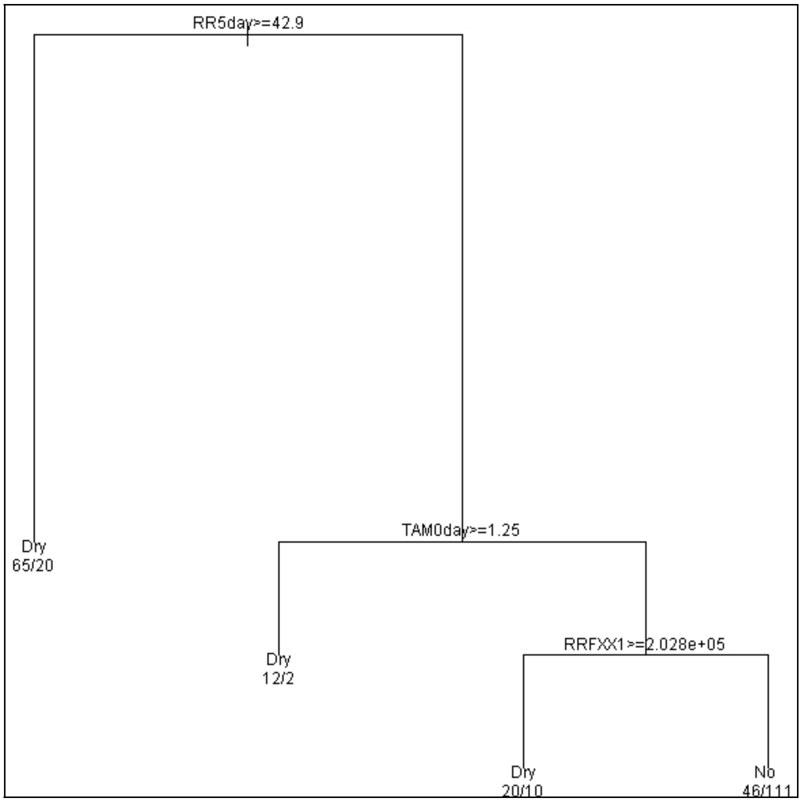
RUN 48


RR2day>=19.4
Dry
58/9
TAM0day>=-3.588
No
43/107
FXM1day>=2.833
Dry
40/20
No
2/7

RUN 49


RR4day>=32.4
FXX0day>=24.1
Dry
66/19
TAN1day>=-8.95
Dry
11/1
TAM1day>=2.075
No
15/67
FXXmax2day>=21.2
Dry
7/0
Dry
12/5
No
32/51

RUN 50


RR5day>=42.9
TAM0day>=1.25
Dry
65/20
RRFXX1>=2.028e+05
Dry
12/2
Dry
20/10
No
46/111