A COMPARISON OF GAP-FILLING METHODS FOR A LONG-TERM EDDY COVARIANCE DATASET FROM A NORTHERN OLD-GROWTH BLACK SPRUCE FOREST

by

Ashley D. Soloway

A Thesis Submitted to the Faculty of Graduate Studies of
The University of Manitoba
in partial fulfilment of the requirements for the degree of

MASTER OF SCIENCE

Department of Soil Science
University of Manitoba
Winnipeg, Manitoba

ABSTRACT

Boreal old-growth forests are key determinants in the global carbon cycle and hold a large amount of the global carbon stock due to their large biomass and peat-rich soils. It is unknown how the role of persistent old-growth forests will be in the carbon cycle in the face of predicted climatic changes. Eddy-covariance measurements are commonly used to quantify carbon exchange between ecosystems, such as forests, and the atmosphere. There are errors associated with these fluxes stemming from measurement equipment and processing steps. Error due to gap-fill method is of particular interest. There have been a few studies determining uncertainty using many forested sites and many gap-fill methods, but fewer have determined the uncertainty among of gap-fill methods for a single site. Here we filled a 15-year eddy covariance, net ecosystem production (NEP) dataset from the Northern Old-Growth Boreal Black Spruce (*Picea mariana*) site located near Thompson, in central Manitoba, Canada using four different gap-fill methods (two non-linear methods, and two variants of lookup table methods). Our objectives were to determine if choice of gap-fill method affected annual NEP and if these errors compounded to even greater differences over the 15-year study period. We also examined the associated flux-partitioning methods and the resulting components of NEP, ecosystem respiration and gross ecosystem production. Average annual NEP values were 4, 22, 44 and 48 g C m^{-2} y^{-1} for each of the four gap-fill methods; where the two largest estimates were significantly greater than the lowest one. Based on the average of all gap-fill methods, this site was a small carbon sink (29 ± 10 g C m^{-2} y^{-1}). Most significant differences in NEP among methods occurred from September to December, but variations during the growing season were responsible for most of the annual differences.

ACKNOWLEDGEMENTS

I'd like to acknowledge the funding sources for the project, including NSERC, NASA, BIOCAP Canada and the Canadian Foundation for Climate and Atmospheric Sciences. The NOBS site was established by Dr. Steve Wofsy (Harvard University), so we would like to thank him, as well as Dr. Allison Dunn (Worcester State University) for the earlier datasets used in the study. I'd also like to thank the NSERC Discovery Grant for stipend funding.

We thank the UBC Biometeorology group (T.A. Black, Z. Nesic and colleagues) and A. Barr (Environment Canada) for the development of many of the MATLAB programming routines. We'd also like to thank the Max Planck Institute of Biogeochemistry (Markus Reichstein and colleagues) for the development of the online gap-fill tool and for related inquiries.

The Northern Old-Growth Black Spruce forest is located on the traditional lands of the Nisichawayasink Cree Nation.

I'd like to thank my project advisor Dr. Brian Amiro for his patience and guidance throughout the project. Thank-you for suggesting this project to me, your support has helped me grow as both a student and a researcher, and has helped me find an area of study I truly enjoy. I'd also like to thank my committee members, Dr. Mario Tenuta and Dr. Tim Papakyriakou for their input into the study and comments to help push the project forward.

A very special thank-you to my friends and family, for their continued support through my lengthy bout as a student and for always encouraging me to strive for my best and to never give up. I'd be nowhere without them in my life

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

1.1 Overview

Over the past century, atmospheric CO_2 concentrations have been rising rapidly (NOAA 2015), mainly due to increased anthropogenic emissions. Forests play a huge role in the carbon cycle, holding a large proportion of global carbon stores (Pan et al. 2011). Northern forests have enormous capacity to sequester carbon and keep it relatively stable for long periods of time without large-scale disturbance (Amiro et al. 2010, Kurz et al. 2013). Their ability to take up CO_2 from the atmosphere makes them an attractive option to offset increasing emissions. The problem is that it is unknown how persistent they can be in holding carbon for extended periods of time with possible climate shifts. Northern forests are particularly vulnerable to predicted changes in climate due to their location (Soja et al. 2007). Recent trends, such as increasing temperatures and atmospheric CO_2, have increased the need for proper carbon accounting and appropriate, more effective mitigation strategies. Better understanding of old-growth forest carbon dynamics is needed for predicting response to further increases in CO_2 concentrations, increasing temperatures, changing precipitation regimes and possible increases in natural disturbances. Flux measurements can be used to estimate carbon exchange between a forest and the atmosphere, but there is uncertainty associated with these datasets. The uncertainty stems from instrument and measurement errors, and data processing errors including choice of gap-fill method and quality control measures. A better understanding of the magnitude of uncertainty will greatly increase the applicability and use of flux data.

1.2 The Study of Carbon Fluxes

The study of carbon flux measurements has been ongoing and understanding has increased greatly since the implementation of long-term flux towers over two decades ago. Carbon flux measurements are important in characterizing the global carbon cycle and fully understanding the movement of carbon through ecosystems and the atmosphere. Carbon fluxes can give insight into carbon sequestration, greenhouse gas emissions, release of carbon to the atmosphere and ecosystem behaviour. Flux studies encompass many techniques such as chamber measurements, thermal stability techniques, mass balance methods, and micrometeorological techniques, such as eddy covariance or eddy accumulation (Denmead et al. 2008) as well as flux gradient methods (Glenn et al. 2010), and energy balance methods (Denmead et al. 2008). Continuous flux measurements are often used for long-term ecosystem monitoring of net gaseous carbon emissions such as CO_2 and CH_4, along with other trace gases and non-carbon containing compounds such as N_2O and H_2O. The information can then be used to characterize and optimize emission mitigation, and determine carbon, energy or water balances.

Flux measurements facilitate long-term carbon studies where ecosystems would otherwise be difficult to measure due to landscape heterogeneity and overall size. Most other carbon measurement techniques are much more involved, requiring multiple sampling trips and are not robust enough to detect subtle annual changes (Campbell et al. 2000). For example, physical soil carbon measurements must be done every 5-10 years with samplings at multiple depths (Campbell et al. 2000), and still, small changes in soil carbon may not be detected. The issue of consistent and/or appropriate sampling times may also arise. Growth and yield information are used in conjunction with known management, statistics on natural disturbances, and changes in carbon stocks to obtain forest inventory data that are as accurate as possible (Kurz et al. 2013). In this way we can estimate carbon content in forests. These studies also require multiple trips over time in order to characterize the progression of tree size, composition and changes due to natural disturbance and samples may only represent a few choice areas in a large ecosystem.

Boreal forests are key determinants in the global carbon cycle due to their capacity to retain large amounts of carbon in biomass and soils (Tarnocai et al. 2009, Kurz et al. 2013). On a global scale, boreal forests have removed just over 20% of all forest-removed anthropogenic fossil fuel emissions since the 1990s (Pan et al. 2011). The boreal forest spans a vast area in the northern hemisphere including North America and Northern Europe. Canada contributes nearly a quarter of all boreal forests globally (Brandt et al. 2009). The location of the boreal biome makes the forests within particularly prone to changes in response to our rapidly changing climate (Soja et al. 2007). These forests are subject to the possibility of changes in forest and carbon dynamics, species shifts and redistributions, treeline shifts, pests and disease, and possible increases in productivity (Kurz et al. 2013) in response to increasing temperatures (Briffa et al. 2008) and atmospheric CO_2 concentrations (Hickler et al. 2008).

The eddy covariance technique is the most common method for monitoring CO_2 exchange between large ecosystems, like forests, and the atmosphere (Baldocchi et al. 2008). Eddy-covariance towers are set up extending above the forest canopy and can be left with minimal intervention. Flux data can be measured in both managed and unmanaged forests and studies can span a long period of time to capture the slowly changing carbon dynamics. There are numerous flux tower sites that have now been in existence for over a decade, for example, the Harvard forest flux tower in Massachusetts, USA (http://fluxnet.ornl.gov/site/886). However, there has yet to be a tower in place for the entirety of boreal forest succession. In studying forest chronosequences, we can obtain a picture of the different flux characteristics of forests of different ages. The changes in boreal forest carbon balance through succession are well understood. Amiro et al. (2010) studied a wide range of forest and disturbance types finding that following disturbance there was always a net loss in carbon in the youngest forests. In boreal forests specifically, maximum carbon loss reached approximately 200 g C m^{-2} y^{-1} (Amiro et al. 2010). Carbon loss in forests from warmer climates (temperate or tropical) was 5-6 times greater than that in boreal forests. All forests became carbon sinks after approximately 20 years; boreal forests remained constant at this rate of carbon gain over time with approximately 210 g C m^{-2} y^{-1} (Amiro et al. 2010). In a separate chronosequence study, boreal forests were a carbon source for approximately nine years when, depending on stand type, a shift from source to sink would occur (Coursolle et al. 2012). A boreal black spruce forest (*Picea mariana*) was a sink after only 10 years, a jack pine (*Pinus banksiana*) stand became a sink by 14 years, and a Douglas fir (*Pseudotsuga menziesii*) stand became a sink after 18 years (Coursolle et al. 2012). The initial carbon losses by these forests were offset by 19, 34 and 47 years respectively, of regrowth (Coursolle et al. 2012). This is similar to other North American boreal chronosequences (Zha et al. 2009, Goulden et al. 2011).

Mkhabela et al. (2008) found regardless of disturbance type, youngest boreal forest sites were always carbon sources, and those aged approximately 20 years were always the largest carbon sinks. After fire disturbance, it was hypothesized that forests go through four phases (Mkhabela et al. 2008). Immediately after disturbance, the forest is a net carbon source, becomes a carbon sink, and then potentially enters a second state of carbon loss due to the decomposition of leftover woody debris (depending on disturbance type), followed by forest maturity characterized by carbon neutrality or by a small carbon sink (Mkhabela et al. 2008). The role disturbance plays on growing season length appears to explain some of the shift in forests from source to sink (Coursolle et al. 2012). Growing season length in young, recently disturbed forests were anywhere from 9 to 109 days shorter than the oldest stands in the study (Coursolle et al. 2012). Around forest age 15-20 years, gross ecosystem production (GEP) growing season length increased sufficiently to match those of older stands; incidentally, this is also the range in which most forests shift from source to sink (Coursolle et al. 2012). Disturbances that occur in forested ecosystems range from natural, such as lightning strike, windthrow, fire and insect infestation; to human induced, such as harvesting and prescribed burn or clear-cutting for land-use change. Fire is the driving disturbance in boreal forest carbon balance (Amiro et al. 2006). Response to disturbance will also depend heavily on how forest dynamics may be altered with climate change.

The role of old-growth forests in the global carbon cycle has been under debate due to the shift in understanding of old-growth forest carbon dynamics. The classical view of old-growth forests was that their ability to sequester carbon diminished after a certain age, and would eventually become a net carbon source (Odum 1969). This theory has been supported mostly by observations from even-aged tree plantations (Gower et al. 1996). But in recent analyses this has not held true (Luyssaert et al. 2008). In a world-wide analysis of old-growth forests, it was far more common for forests to be net carbon sinks or carbon neutral (Luyssaert et al. 2008). Although decreases in net ecosystem production (NEP) in old-growth boreal forests (154 years) have been observed, there were still moderate amounts of carbon gained. This decrease was attributed to loss of live biomass carbon and a rise in respiration due to increased decomposition on the forest floor (Luyssaert et al. 2008). It has also been observed that old-growth forest soils can continue to accumulate carbon at a rate of approximately 0.035% per year in southern China (Zhou et al. 2006). Due to the longevity of old-growth forests, the fact that boreal forests are the largest carbon sink in the northern hemisphere (Pan et al. 2011) and the uncertainty with which their dynamics will change with climate change, it is important to monitor over time to better implement plans for mitigation of carbon emissions and increasing our predictive power as to how carbon sinks may adapt to new climates (Kurz et al. 2013).

1.3 Flux Tower Measurements

FLUXNET is a network with hundreds of flux tower stations across the globe measuring energy, trace gases, carbon and water fluxes as well as meteorological data over all sorts of landscapes, agricultural sites, water bodies, wetlands and forest types (http://fluxnet.ornl.gov/). The tower instruments give an estimate of carbon exchange on a local scale, usually around 1 km in diameter. FLUXNET facilitates cooperation and data sharing among researchers and provides access to datasets on a site by site basis.

Measuring CO_2 flux gives insight into an ecosystem's net ecosystem exchange (NEE). When we assume there is no loss of soil dissolved carbon and other carbon containing gases, the following is true:

$$NEP = -NEE \qquad \text{(Eq. 1.1)}$$

We can also partition NEP into its components of GEP and ecosystem respiration (R):

$$NEP = GEP - R \qquad \text{(Eq. 1.2)}$$

This convention, mostly used by ecologists, defines positive NEP as net carbon flux towards the forest (carbon gain). Positive GEP corresponds to carbon uptake and positive R corresponds to carbon lost to the atmosphere.

Flux datasets are not only useful in carbon balance studies (e.g. Dunn et al. 2007), but also for energy budgets during growing seasons (e.g. Lei et al. 2010) or ecosystem water balances (e.g. Williams et al. 2012). Comparisons with flux data can be made to assess the variability of fluxes temporally and spatially, between or within sites and ecosystems, and based on meteorological conditions (Baldocchi et al. 2001). Sites can run for weeks or months or for many years in order to get a good picture of the flux characteristics (e.g. NEE, NEP, GEP and R) over time. Flux data are often used to validate ecosystem models (Li et al. 2004) and can also be used to provide broader landscape analyses by upscaling measurements (Xiao et al. 2012).

The Northern Old-Growth Black Spruce (*Picea mariana*) (NOBS) forest located in central Manitoba, Canada, near Thompson at the northern limit of the boreal forest in the Canadian Shield, is a former FLUXNET tower site. This site is unique as it is one of few in the North American boreal zone, and also ran for nearly 15 years. It was established in 1994 as part of NASA's boreal ecosystem atmosphere study (BOREAS) as an environmental measurement system with the objective to monitor atmosphere-biosphere interactions for a typical boreal old-growth boreal forest (Sellers et al. 1995). A wide variety of studies have been performed using the eddy covariance data from NOBS ranging from model validation (Turner et al. 2006, Heinsch et al. 2006, Bonan et al. 2011, Stoy et al. 2014, Hilton et al. 2014), model creation (Horn and Schulz 2011, Fu et al. 2014), model comparison (Raczka et al. 2013), carbon flux analysis within the forest (Rocha et al. 2006, Dunn et al. 2007) and has been used among North American forests encompassing effects of environmental changes (Mahecha et al. 2010, Schwalm et al. 2010), chronosequences (Goulden et al. 2006, Amiro et al. 2010, Goulden et al. 2011, Coursolle et al. 2012) and across longitudinal gradients (Coursolle et al. 2006, Bergeron et al. 2007). Most of these examples were published after the NOBS flux tower site was decommissioned in 2008 and there are numerous other studies published before those mentioned. The amount NOBS flux data have been used, even after its decommissioning, indicates the importance of long-term sites and the validation of flux data.

Eddy-covariance data are subject to error in multiple forms. The total uncertainty of eddy-covariance measurements includes error in the actual measurement, the calculation of the flux, uncertainty in screening of fluxes (i.e. outlier removal), selection of wind velocity threshold, instrumental error (e.g. instrumental self-heating correction), and error attributed to selection of gap-fill method (Elbers et al. 2011). It is difficult to quantify error and many researchers do not report results with uncertainty estimates (Elbers et al. 2011). To compound this issue, some researchers will only report either uncertainty of one type or a selection of a few types. In a study of a Scots pine (*Pinus sylvestris*) forest in the Netherlands, it was found that the largest contributions to total uncertainty were from error in flux calculation and error due to selection of wind velocity threshold, while lowest contribution was from self-heating correction of instrumentation (Elbers et al. 2011). Similarly, correcting for wind velocity threshold shifted annual NEP on average -77 g C m^{-2} y^{-1} but at times as much as -185 g C m^{-2} y^{-1} (Falge et al. 2001). Total uncertainty averaged approximately ± 32 g C m^{-2} y^{-1} (Elbers et al. 2011) which was close to the uncertainty reported by Baldocchi (2003) of ± 50 g C m^{-2} y^{-1} determined from multiple forest types. This appears to be on the lower end of total uncertainty estimates; Rannik et al. (2006) estimated total uncertainty to be approximately ± 80 g C m^{-2} y^{-1} at a Scots pine forest while in a loblolly pine (*Pinus taeda*) plantation, uncertainty in a given year could range from ± 64 to 110 g C m^{-2} y^{-1} (Oren et al. 2006). The reason for small uncertainty in the study by Elbers et al. (2011) was thought to be due to low incidence of long gaps which created a rather small gap-filling error. Long gaps create larger error when filled (Richardson and Hollinger 2007). In a study determining random uncertainty in flux datasets, it was determined that gap-fill method was the largest contributor to overall random uncertainty (Dragoni et al. 2007).

Long-term eddy flux measurements allow basic monitoring over time and provide an excellent opportunity to observe ecosystem responses over a large array of meteorological states. There will inevitably be times when measurements will not be taken due to system malfunction or power failure which create larger data gaps, or issues of unsuitable weather conditions like rain, low wind turbulence, or snow covering sensors, which lead to smaller gaps (Dragoni et al. 2007). Data gaps may be up to 70% of a dataset on an annual basis (Moffat et al. 2007). Flux datasets also undergo quality control in order to ensure that all accepted data are within realistic bounds and represent true carbon or energy flux. This entails excluding measurements taken in low turbulence conditions, which are usually adjusted based on site conditions including surface roughness (Goulden et al. 1997, Papale et al. 2006) and may include directional exclusions or flux limits. Using both remaining data and associated meteorological data, gap-fill methods have been developed in order to give a best estimate of the missing flux at that time point. These methods differ greatly depending on algorithms used, whether it is lookup tables, mean diurnal variation (Falge et al. 2001), regressions (Barr et al. 2004, Desai et al. 2005, Reichstein et al. 2005) or neural networks (Papale and Valentini 2003, Braswell et al. 2005).

Gap-fill methods can often partition NEP into component fluxes GEP and R. This is useful to better understand plant and soil functional responses to changes in meteorological variables and how these responses affect the observed NEP (Eq. 1.2). This is one of the only ways to get this type of information since GEP and R encompass many difference responses for many different organism types; autotrophic for both GEP and R and heterotrophic for R. As with gap-filling, there is no standard method for flux-partitioning (Reichstein et al. 2005, Stoy et al. 2006). Many methods rely on the fact that there is no GEP in nighttime NEP measurements (Reichstein et al. 2005) which are riddled with data gaps due to low turbulence, and requires the assumption that GEP and R are tightly linked with environmental variables that drive the responses (Law et al. 2002). Desai et al. (2008) found using nighttime data and short-term temperature sensitivity (Reichstein et al. 2005) yielded the best GEP and R estimates in comparison to synthetic data produced by an ecosystem model. Contrarily, Stoy et al. (2006) found more complex methods using daytime measurements to derive R yield the most accurate results. Neural networks have been suggested as a superior method since they have less underlying assumptions, and can be tailored to different ecosystems (Desai et al. 2008). Flux-partitioning methods compared in Desai et al. (2008) agreed on the ranking of sites relative to one another as far as which were greatest carbon sinks and which were sources, as well as their seasonal trends; but in general it has been found that in order to compare GEP and R between sites it is important to use the same flux-partitioning method (Desai et al. 2008).

1.4 Gaps in Knowledge

As previously mentioned, total uncertainty is made up of many parts, each with challenges in quantification (Elbers et al. 2011). Specifically, we are interested here in gap-fill uncertainty. There have been few studies addressing gap-fill error specifically (Falge et al. 2001, Ooba et al. 2006, Moffat et al. 2007) and on an annual basis, it has been determined as important to overall uncertainty.

A large-scale study comparing multiple gap-fill methods among many different forested sites determined that gap-fill methods fell within a range of ± 25 g C m^{-2} y^{-1} (Moffat et al. 2007). In another study, gap-fill error fell within a range of ± 7 g C m^{-2} y^{-1}

(Elbers et al. 2011). This disagreement was attributed to a low occurrence of large gaps in the dataset used in Elbers et al. (2011). A study by Falge et al. (2001) reported that values of NEE differed between gap fill methods by -45 to + 200 g C m^{-2} y^{-1} (mean diurnal variation compared to nonlinear regression) and by -30 to +150 g C m^{-2} y^{-1} between a look-up method and nonlinear regression, when filling flux datasets of multiple forested sites. Overall it was estimated choice of gap-fill method added annual error in the range of 0.83 g C m^{-2} per percentage of year filled for daytime measurements (Falge et al. 2001). This is greater than compared to nighttime error (± 0.52 g C m^{-2} per % of year filled) (Falge et al. 2001). That means in a dataset with 50% gaps, error could be as much as ± 41.5 g C m^{-2} y^{-1} for filling daytime measurements and an additional ± 26 g C m^{-2} y^{-1} for filling nighttime data. Thus far, studies that have focused on gap-fill method comparison within a site have had less success than studies encompassing more sites due to shorter measurement periods and the use of managed ecosystems. Measuring CO_2 flux over arable lands in the Netherlands, it was determined that gap-filling depended greatly on season and environmental conditions with greater accuracy of estimates in winter measurements (Dragomir et al. 2012). It was concluded the artificial neural network was the most accurate filling technique and minimized the seasonal effects (Dragomir et al. 2012). It should be noted, however, that winter measurements are proportionally smaller than summer CO_2 fluxes in any ecosystem and are usually near zero; especially in colder climates. That being said, numerous other gap-filling studies have elected neural networks as a superior gap-filling strategy (Falge et al. 2001, Ooba et al. 2006, Papale et al. 2006, Moffat et al. 2007, Desai et al. 2008). Another within-site study analyzing gap-fill methods looked at flux measurements from a Japanese larch (*Larix kaempferi*) plantation but only encompassed the growing season of one year (Ooba et al. 2006).

Daily NEP values were similar when compared between neural networks (-3.48 g C m^{-2} d^{-1}) and those filled with empirical equations (-2.76 g C m^{-2} d^{-1}) (Ooba et al. 2006). Whether comparing among or within sites, short time scales and homogenous ecosystems may not reflect long-term trends (Reichstein et al. 2015) or the response of natural ecosystems.

There are also errors associated with flux partitioning; which involves partitioning NEP into its component fluxes R and GEP. Luckily, their errors are offest so that by modelling them to use for gap-filling NEP, there is much smaller amount of error in NEP (Moffat et al. 2007). Hagen et al. (2006) partitioned NEP measured from boreal transition forests using multiple methods; GEP estimates varied by more than 100 g C m^{-2} y^{-1}. Similarly, the inter-quartile range (75%) of methods were all within approximately 100 g C m^{-2} y^{-1} in a study by Desai et al. (2008). These errors are large enough to disguise true differences among sites with regards to GEP and R. Due to all quality control measures, Papale et al. (2006) estimate that less than 100 g C m^{-2} y^{-1} is added to NEP, but this will translate to larger error in partitioned GEP and R. Assessing all sources of error in GEP and R, Desai et al. (2008) estimate error may approach 25%. This would actually limit the use of flux-partitioned data, but only if compared partitioning components were obtained by different methods (Desai et al. 2008).

Understanding the error associated with gap-filling data in all aspects is important due to the crucial role flux data play in designating forests as sources or sinks. This can have great effects on management decisions as to how forests can be accounted for in carbon mitigation plans. Better characterizing eddy-covariance gap-fill error will allow better advancement in data usage and estimates will become more reliable and accurate.

This could help by giving better estimates in ecosystem monitoring during periods of high data gaps and could help improve accuracy in a newer application for eddy-covariance data where flux is up-scaled in an attempt to estimate exchange over larger areas of land (Xiao et al. 2012). However, most of the issues in up-scaling have to do with insufficient coverage of land types proportional to the area being estimated (Xiao et al. 2012). Having a better idea of the magnitude of gap-fill error allows eddy covariance to be more applicable in global carbon budgets and better in predicting carbon's response to changing climate, disturbance and other forcing variables. It has been suggested by many that in order to obtain better comparisons between and among sites and ecosystem types, as well as to facilitate synthesis activities, that eddy covariance measurements would benefit from standardized processing routines (Barr et al. 2002, Papale et al. 2006, Moffat et al. 2007, Desai et al. 2008). When data are measured and processed differently, comparisons are made unreliable because errors added to each dataset are completely different from one another. It would be difficult to tell if observed differences are caused by differences between sites or whether differences are attributable to differences in processing. Being able to quantify gap-fill uncertainty will contribute to alleviating this problem and may allow better flux characterization.

1.5 Study Objectives

The main objective of this study is to determine the effect of gap-filling on a long-term eddy-covariance dataset. We filled one of the longest available boreal forest datasets for the NOBS forest that spans 15 years with four different gap-fill methods to assess the effect on annual NEP. In studies so far, it seems as though on an annual basis, choice of gap-filling method does not affect eddy-covariance results (Moffat et al. 2007). It is unknown however, if these errors compound over time in a long-term dataset. We would then like to determine (if there are differences among methods) if flux partitioning can give any indication of the reasons for differences or if it is simply dependent on the equations with which gaps are filled. We hypothesize that annually, the NOBS carbon balance will not differ based on gap-fill method.

This document consists of this overall introduction, followed by a manuscript detailing the experimental procedures and study results, and ends with an overall synthesis for the project.

1.6 References

Amiro, B. D., Barr, A. G., Barr, J. G., Black, T. A., Bracho, R., Brown, M., Chen, J., Clark, K. L., Davis, K. J., Desai, A. R., Dore, S., Engel, V., Fuentes, J. D., Goldstein, A. H., Goulden, M. L., Kolb, T. E., Lavigne, M. B., Law, B. E., Margolis, H. A., Martin, T., McCaughey, J. H., Misson, L., Montes-Helu, M., Noormets, A., Randerson, J. T., Starr, G. and Xiao, J. 2010. Ecosystem carbon dioxide fluxes after disturbance in forests of North America. J. Geophys. Res-Biogeo. **115:** G00K02.

Amiro, B. D., Orchansky, A. L., Barr, A. G., Black, T. A., Chambers, S. D., Chapin, F. S.,III, Goulden, M. L., Litvakg, M., Liu, H. P., McCaughey, J. H., McMillan, A. and Randerson, J. T. 2006. The effect of post-fire stand age on the boreal forest energy balance. Agric. For. Meteorol. **140:** 41-50.

Baldocchi, D. D. 2003. Assessing the eddy covariance technique for evaluating carbon dioxide exchange rates of ecosystems: past, present and future. Global Change Biol. **9:** 479-492.

Baldocchi, D. D. and Wilson, K. B. 2001. Modeling CO_2 and water vapor exchange of a temperate broadleaved forest across hourly to decadal time scales. Ecol. Model. **142:** 155-184.

Baldocchi, D. 2008. Breathing of the terrestrial biosphere: lessons learned from a global network of carbon dioxide flux measurement systems. Aust. J. Bot. **56:** 1-26.

Baldocchi, D., Sturtevant, C. and Fluxnet Contributors. 2015. Does day and night sampling reduce spurious correlation between canopy photosynthesis and ecosystem respiration? Agric. For. Meteorol. **207:** 117-126.

Barr, A. G., Black, T. A., Hogg, E. H., Kljun, N., Morgenstern, K. and Nesic, Z. 2004. Inter-annual variability in the leaf area index of a boreal aspen-hazelnut forest in relation to net ecosystem production. Agric. For. Meteorol. **126:** 237-255.

Barr, A. G., Griffis, T. J., Black, T. A., Lee, X., Staebler, R. M., Fuentes, J. D., Chen, Z. and Morgenstern, K. 2002. Comparing the carbon budgets of boreal and temperate deciduous forest stands. Can. J. Forest Res. **32:** 813-822.

Bergeron, O., Margolis, H. A., Black, T. A., Coursolle, C., Dunn, A. L., Barr, A. G. and Wofsy, S. C. 2007. Comparison of carbon dioxide fluxes over three boreal black spruce forests in Canada. Global Change Biol. **13:** 89-107.

Bonan, G. B., Lawrence, P. J., Oleson, K. W., Levis, S., Jung, M., Reichstein, M., Lawrence, D. M. and Swenson, S. C. 2011. Improving canopy processes in the

Community Land Model version 4 (CLM4) using global flux fields empirically inferred from FLUXNET data. J. Geophys. Res-Biogeo. **116:** G02014.

Brandt, J. P. 2009. The extent of the North American boreal zone. Env. Rev. **17:** 101-161.

Braswell, B. H., Sacks, W. J., Linder, E. and Schimel, D. S. 2005. Estimating diurnal to annual ecosystem parameters by synthesis of a carbon flux model with eddy covariance net ecosystem exchange observations. Global Change Biol. **11:** 335-355.

Briffa, K. R., Shishov, V. V., Melvin, T. M., Vaganov, E. A., Grudd, H., Hantemirov, R. M., Eronen, M. and Naurzbaev, M. M. 2008. Trends in recent temperature and radial tree growth spanning 2000 years across northwest Eurasia. Philos. T. Roy. Soc. B. **363:** 2271-2284.

Campbell, C. A., Zentner, R. P., Liang, B. C., Roloff, G., Gregorich, E. C. and Blomert, B. 2000. Organic C accumulation in soil over 30 years in semiarid southwestern Saskatchewan - Effect of crop rotations and fertilizers. Can. J. Soil Sci. **80:** 179-192.

Coursolle, C., Margolis, H. A., Barr, A. G., Black, T. A., Amiro, B. D., McCaughey, J. H., Flanagan, L. B., Lafleur, P. M., Roulet, N. T., Bourque, C. P. A., Arain, M. A., Wofsy, S. C., Dunn, A., Morgenstern, K., Orchansky, A. L., Bernier, P. Y., Chen, J. M., Kidston, J., Saigusa, N. and Hedstrom, N. 2006. Late-summer carbon fluxes from Canadian forests and peatlands along an east-west continental transect. Can. J. Forest Res. **36:** 783-800.

Coursolle, C., Margolis, H. A., Giasson, M. A., Bernier, P. Y., Amiro, B. D., Arain, M. A., Barr, A. G., Black, T. A., Goulden, M. L., McCaughey, J. H., Chen, J. M., Dunn, A. L., Grant, R. F. and Lafleur, P. M. 2012. Influence of stand age on the magnitude and seasonality of carbon fluxes in Canadian forests. Agric. For. Meteorol. **165:** 136-148.

Denmead, O. T. 2008. Approaches to measuring fluxes of methane and nitrous oxide between landscapes and the atmosphere. Plant Soil **309:** 5-24.

Desai, A. R., Bolstad, P. V., Cook, B. D., Davis, K. J. and Carey, E. V. 2005. Comparing net ecosystem exchange of carbon dioxide between an old-growth and mature forest in the upper Midwest, USA. Agric. For. Meteorol. **128:** 33-55.

Desai, A. R., Richardson, A. D., Moffat, A. M., Kattge, J., Hollinger, D. Y., Barr, A., Falge, E., Noormets, A., Papale, D., Reichstein, M. and Stauch, V. J. 2008. Cross-site evaluation of eddy covariance GPP and RE decomposition techniques. Agric. For. Meteorol. **148:** 821-838.

Dragomir, C. M., Klaassen, W., Voiculescu, M., Georgescu, L. P. and van der Laan, S. 2012. Estimating annual CO_2 flux for Lutjewad Station using three different gap-filling techniques. Scientific World Journal 1-10.

Dragoni, D., Schmid, H. P., Grimmond, C. S. B. and Loescher, H. W. 2007. Uncertainty of annual net ecosystem productivity estimated using eddy covariance flux measurements. J. Geophys. Res.-Atmos. **112:** D17102.

Dunn, A. L., Barford, C. C., Wofsy, S. C., Goulden, M. L. and Daube, B. C. 2007. A long-term record of carbon exchange in a boreal black spruce forest: means, responses to interannual variability, and decadal trends. Global Change Biol. **13:** 577-590.

Elbers, J. A., Jacobs, C. M. J., Kruijt, B., Jans, W. W. P. and Moors, E. J. 2011. Assessing the uncertainty of estimated annual totals of net ecosystem productivity: A practical approach applied to a mid latitude temperate pine forest. Agric. For. Meteorol. **151:** 1823-1830.

Falge, E., Baldocchi, D., Olson, R., Anthoni, P., Aubinet, M., Bernhofer, C., Burba, G., Ceulemans, R., Clement, R., Dolman, H., Granier, A., Gross, P., Grunwald, T., Hollinger, D., Jensen, N. O., Katul, G., Keronen, P., Kowalski, A., Lai, C. T., Law, B. E., Meyers, T., Moncrieff, H., Moors, E., Munger, J. W., Pilegaard, K., Rannik, U., Rebmann, C., Suyker, A., Tenhunen, J., Tu, K., Verma, S., Vesala, T., Wilson, K. and Wofsy, S. 2001. Gap filling strategies for defensible annual sums of net ecosystem exchange. Agric. For. Meteorol. **107:** 43-69.

Fu, D., Chen, B., Zhang, H., Wang, J., Black, T. A., Amiro, B. D., Bohrer, G., Bolstad, P., Coulter, R., Rahman, A. F., Dunn, A., McCaughey, J. H., Meyers, T. and Verma, S. 2014. Estimating landscape net ecosystem exchange at high spatial-temporal resolution based on Landsat data, an improved upscaling model framework, and eddy covariance flux measurements. Remote Sens. Environ. **141:** 90-104.

Glenn, A. J., Amiro, B. D., Tenuta, M., Stewart, S. E. and Wagner-Riddle, C. 2010. Carbon dioxide exchange in a northern Prairie cropping system over three years. Agric. For. Meteorol. **150:** 908-918.

Goulden, M. L., Daube, B. C., Fan, S. M., Sutton, D. J., Bazzaz, A., Munger, J. W. and Wofsy, S. C. 1997. Physiological responses of a black spruce forest to weather. J. Geophys. Res.-Atmos. **102:** 28987-28996.

Goulden, M. L., McMillan, A. M. S., Winston, G. C., Rocha, A. V., Manies, K. L., Harden, J. W. and Bond-Lamberty, B. P. 2011. Patterns of NPP, GPP, respiration, and NEP during boreal forest succession. Global Change Biol. **17:** 855-871.

Goulden, M. L., Winston, G. C., McMillan, A. M. S., Litvak, M. E., Read, E. L., Rocha, A. V. and Elliot, J. R. 2006. An eddy covariance mesonet to measure the effect of forest age on land-atmosphere exchange. Global Change Biol. **12:** 2146-2162.

Gower, S. T., McMurtrie, R. E. and Murty, D. 1996. Aboveground net primary production decline with stand age: Potential causes. Trends Ecol. Evol. **11:** 378-382.

Hagen, S. C., Braswell, B. H., Linder, E., Frolking, S., Richardson, A. D. and Hollinger, D. Y. 2006. Statistical uncertainty of eddy flux-based estimates of gross ecosystem carbon exchange at Howland Forest, Maine. J. Geophys. Res.-Atmos. **111:** D08S03.

Heinsch, F. A., Zhao, M., Running, S. W., Kimball, J. S., Nemani, R. R., Davis, K. J., Bolstad, P. V., Cook, B. D., Desai, A. R., Ricciuto, D. M., Law, B. E., Oechel, W. C., Kwon, H., Luo, H., Wofsy, S. C., Dunn, A. L., Munger, J. W., Baldocchi, D. D., Xu, L., Hollinger, D. Y., Richardson, A. D., Stoy, P. C., Siqueira, M. B. S., Monson, R. K., Burns, S. P. and Flanagan, L. B. 2006. Evaluation of remote sensing based terrestrial productivity from MODIS using regional tower eddy flux network observations. IEEE Trans. Geosci. Remote Sens. **44:** 1908-1925.

Hickler, T., Smith, B., Prentice, I. C., Mjofors, K., Miller, P., Arneth, A. and Sykes, M. T. 2008. CO_2 fertilization in temperate FACE experiments not representative of boreal and tropical forests. Global Change Biol. **14:** 1531-1542.

Hilton, T. W., Davis, K. J. and Keller, K. 2014. Evaluating terrestrial CO_2 flux diagnoses and uncertainties from a simple land surface model and its residuals. Biogeosciences **11:** 217-235.

Horn, J. E. and Schulz, K. 2011. Identification of a general light use efficiency model for gross primary production. Biogeosciences **8:** 999-1021.

Kurz, W. A., Shaw, C. H., Boisvenue, C., Stinson, G., Metsaranta, J., Leckie, D., Dyk, A., Smyth, C. and Neilson, E. T. 2013. Carbon in Canada's boreal forest - A synthesis. Env. Rev. **21:** 260-292.

Law, B. E., Falge, E., Gu, L., Baldocchi, D. D., Bakwin, P., Berbigier, P., Davis, K., Dolman, A. J., Falk, M., Fuentes, J. D., Goldstein, A., Granier, A., Grelle, A., Hollinger, D., Janssens, I. A., Jarvis, P., Jensen, N. O., Katul, G., Mahli, Y., Matteucci, G., Meyers, T., Monson, R., Munger, W., Oechel, W., Olson, R., Pilegaard, K., Paw U, K. T., Thorgeirsson, H., Valentini, R., Verma, S., Vesala, T., Wilson, K. and Wofsy, S. 2002. Environmental controls over carbon dioxide and water vapor exchange of terrestrial vegetation. Agric. For. Meteorol. **113:** 97-120.

Lei, H. and Yang, D. 2010. Interannual and seasonal variability in evapotranspiration and energy partitioning over an irrigated cropland in the North China Plain. Agric. For. Meteorol. **150:** 581-589.

Li, T., Grant, R. F. and Flanagan, L. B. 2004. Climate impact on net ecosystem productivity of a semi-arid natural grassland: modeling and measurement. Agric. For. Meteorol. **126:** 99-116.

Luyssaert, S., Schulze, E. D., Boerner, A., Knohl, A., Hessenmoeller, D., Law, B. E., Ciais, P. and Grace, J. 2008. Old-growth forests as global carbon sinks. Nature **455:** 213-215.

Mahecha, M. D., Reichstein, M., Jung, M., Seneviratne, S. I., Zaehle, S., Beer, C., Braakhekke, M. C., Carvalhais, N., Lange, H., Le Maire, G. and Moors, E. 2010. Comparing observations and process-based simulations of biosphere-atmosphere exchanges on multiple timescales. J. Geophys. Res-Biogeo. **115:** G02003.

Mkhabela, M. S., Amiro, B. D., Barr, A. G., Black, T. A., Hawthorne, I., Kidston, J., McCaughey, J. H., Orchansky, A. L., Nesic, Z., Sass, A., Shashkov, A. and Zha, T. 2009. Comparison of carbon dynamics and water use efficiency following fire and harvesting in Canadian boreal forests. Agric. For. Meteorol. **149:** 783-794.

Moffat, A. M., Papale, D., Reichstein, M., Hollinger, D. Y., Richardson, A. D., Barr, A. G., Beckstein, C., Braswell, B. H., Churkina, G., Desai, A. R., Falge, E., Gove, J. H., Heimann, M., Hui, D., Jarvis, A. J., Kattge, J., Noormets, A. and Stauch, V. J. 2007. Comprehensive comparison of gap-filling techniques for eddy covariance net carbon fluxes. Agric. For. Meteorol. **147:** 209-232.

NOAA, 2015. Trends in Atmospheric Carbon Dioxide. NOAA/ESRL. [Online] Available: http://www.esrl.noaa.gov/gmd/ccgg/trends/ (Accessed February 12th, 2016).

Odum, E. P. 1969. Strategy of ecosystem development. Science **164:** 262-270.

Ooba, M., Hirano, T., Mogami, J., Hirata, R. and Fujinuma, Y. 2006. Comparisons of gap-filling methods for carbon flux dataset: A combination of a genetic algorithm and an artificial neural network. Ecol. Model. **198:** 473-486.

Oren, R., Hseih, C. I., Stoy, P., Albertson, J., McCarthy, H. R., Harrell, P. and Katul, G. G. 2006. Estimating the uncertainty in annual net ecosystem carbon exchange: spatial variation in turbulent fluxes and sampling errors in eddy-covariance measurements. Global Change Biol. **12:** 883-896.

Pan, Y., Birdsey, R. A., Fang, J., Houghton, R., Kauppi, P. E., Kurz, W. A., Phillips, O. L., Shvidenko, A., Lewis, S. L., Canadell, J. G., Ciais, P., Jackson, R. B., Pacala,

S. W., McGuire, A. D., Piao, S., Rautiainen, A., Sitch, S. and Hayes, D. 2011. A large and persistent carbon sink in the world's forests. Science **333:** 988-993.

Papale, D., Reichstein, M., Aubinet, M., Canfora, E., Bernhofer, C., Kutsch, W., Longdoz, B., Rambal, S., Valentini, R., Vesala, T. and Yakir, D. 2006. Towards a standardized processing of net ecosystem exchange measured with eddy covariance technique: algorithms and uncertainty estimation. Biogeosciences **3:** 571-583.

Papale, D. and Valentini, A. 2003. A new assessment of European forests carbon exchanges by eddy fluxes and artificial neural network spatialization. Global Change Biol. **9:** 525-535.

Raczka, B. M., Davis, K. J., Huntzinger, D., Neilson, R. P., Poulter, B., Richardson, A. D., Xiao, J., Baker, I., Ciais, P., Keenan, T. F., Law, B., Post, W. M., Ricciuto, D., Schaefer, K., Tian, H., Tomelleri, E., Verbeeck, H. and Viovy, N. 2013. Evaluation of continental carbon cycle simulations with North American flux tower observations. Ecol. Monogr. **83:** 531-556.

Rannik, U., Kolari, P., Vesala, T. and Hari, P. 2006. Uncertainties in measurement and modelling of net ecosystem exchange of a forest. Agric. For. Meteorol. **138:** 244-257.

Reichstein, M., Falge, E., Baldocchi, D., Papale, D., Aubinet, M., Berbigier, P., Bernhofer, C., Buchmann, N., Gilmanov, T., Granier, A., Grunwald, T., Havrankova, K., Ilvesniemi, H., Janous, D., Knohl, A., Laurila, T., Lohila, A., Loustau, D., Matteucci, G., Meyers, T., Miglietta, F., Ourcival, J. M., Pumpanen, J., Rambal, S., Rotenberg, E., Sanz, M., Tenhunen, J., Seufert, G., Vaccari, F., Vesala, T., Yakir, D. and Valentini, R. 2005. On the separation of net ecosystem exchange into assimilation and ecosystem respiration: review and improved algorithm. Global Change Biol. **11:** 1424-1439.

Richardson, A. D. and Hollinger, D. Y. 2007. A method to estimate the additional uncertainty in gap-filled NEE resulting from long gaps in the CO_2 flux record. Agric. For. Meteorol. **147:** 199-208.

Rocha, A. V., Goulden, M. L., Dunn, A. L. and Wofsy, S. C. 2006. On linking interannual tree ring variability with observations of whole-forest CO_2 flux. Global Change Biol. **12:** 1378-1389.

Schwalm, C. R., Black, T. A., Morgenstern, K. and Humphreys, E. R. 2007. A method for deriving net primary productivity and component respiratory fluxes from tower-based eddy covariance data: a case study using a 17-year data record from a Douglas-fir chronosequence. Global Change Biol. **13:** 370-385.

Schwalm, C. R., Williams, C. A., Schaefer, K., Anderson, R., Arain, M. A., Baker, I., Barr, A., Black, T. A., Chen, G., Chen, J. M., Ciais, P., Davis, K. J., Desai, A., Dietze, M., Dragoni, D., Fischer, M. L., Flanagan, L. B., Grant, R., Gu, L., Hollinger, D., Izaurralde, R. C., Kucharik, C., Lafleur, P., Law, B. E., Li, L., Li, Z., Liu, S., Lokupitiya, E., Luo, Y., Ma, S., Margolis, H., Matamala, R., McCaughey, H., Monson, R. K., Oechel, W. C., Peng, C., Poulter, B., Price, D. T., Riciutto, D. M., Riley, W., Sahoo, A. K., Sprintsin, M., Sun, J., Tian, H., Tonitto, C., Verbeeck, H. and Verma, S. B. 2010. A model-data intercomparison of CO_2 exchange across North America: Results from the North American Carbon Program site synthesis. J. Geophys. Res-Biogeo. **115:** G00H05.

Sellers, P., Hall, F., Margolis, H., Kelly, B., Baldocchi, D., den Hartog, G., Cihlar, J., Ryan, M. G., Goodison, B., Crill, P., Ranson, K. J., Lettenmaier, D. and Wickland, D. E. 1995. The Boreal Ecosystem-Atmosphere Study (Boreas) - an overview and early results from the 1994 field year. Bull. Am. Meteorol. Soc. **76:** 1549-1577.

Soja, A. J., Tchebakova, N. M., French, N. H. F., Flannigan, M. D., Shugart, H. H., Stocks, B. J., Sukhinin, A. I., Parfenova, E. I., Chapin, F. S.,III and Stackhouse, P. W.,Jr. 2007. Climate-induced boreal forest change: Predictions versus current observations. Global Planet. Change **56:** 274-296.

Stoy, P. C., Katul, G. G., Siqueira, M. B. S., Juang, J., Novick, K. A., Uebelherr, J. M. and Oren, R. 2006. An evaluation of models for partitioning eddy covariance-measured net ecosystem exchange into photosynthesis and respiration. Agric. For. Meteorol. **141:** 2-18.

Stoy, P. C., Trowbridge, A. M. and Bauerle, W. L. 2014. Controls on seasonal patterns of maximum ecosystem carbon uptake and canopy-scale photosynthetic light response: contributions from both temperature and photoperiod. Photosynthesis Res. **119:** 49-64.

Tarnocai, C., Canadell, J. G., Schuur, E. A. G., Kuhry, P., Mazhitova, G. and Zimov, S. 2009. Soil organic carbon pools in the northern circumpolar permafrost region. Global Biogeochem. Cycles **23:** GB2023.

Turner, D. P., Ritts, W. D., Zhao, M., Kurc, S. A., Dunn, A. L., Wofsy, S. C., Small, E. E. and Running, S. W. 2006. Assessing interannual variation in MODIS-based estimates of gross primary production. IEEE Trans. Geosci. Remote Sens. **44:** 1899-1907.

Williams, C. A., Reichstein, M., Buchmann, N., Baldocchi, D., Beer, C., Schwalm, C., Wohlfahrt, G., Hasler, N., Bernhofer, C., Foken, T., Papale, D., Schymanski, S. and Schaefer, K. 2012. Climate and vegetation controls on the surface water balance:

Synthesis of evapotranspiration measured across a global network of flux towers. Water Resour. Res. **48:** W06523.

Xiao, J., Chen, J., Davis, K. J. and Reichstein, M. 2012. Advances in upscaling of eddy covariance measurements of carbon and water fluxes. J. Geophys. Res-Biogeo. **117:** G00J01.

Zha, T., Barr, A. G., Black, T. A., McCaughey, J. H., Bhatti, J., Hawthorne, I., Krishnan, P., Kidston, J., Saigusa, N., Shashkov, A. and Nesic, Z. 2009. Carbon sequestration in boreal jack pine stands following harvesting. Global Change Biol. **15:** 1475-1487.

Zhou, G., Liu, S., Li, Z., Zhang, D., Tang, X., Zhou, C., Yan, J. and Mo, J. 2006. Old-growth forests can accumulate carbon in soils. Science **314:** 1417-1417.

2. A COMPARISON OF GAP-FILLING METHODS FOR A LONG-TERM EDDY COVARIANCE DATASET FROM A NORTHERN OLD-GROWTH BLACK SPRUCE FOREST

2.1 Abstract

Old-growth boreal forests play a critical role in the carbon balance of the northern hemisphere. Long-term flux datasets provide an excellent opportunity to study how boreal forests will respond to changing climate, disturbances and environmental pressures. Trace gases, including CO_2, can be measured by flux towers using the eddy-covariance technique and can give insight into the net ecosystem production (NEP) of the forest. Flux data are subject to data gaps due to quality-control measures and unsuitable meteorological conditions that make measurements unreliable. Gaps in a flux dataset can range from 30% to 70% on an annual basis. Gap-fill methods have been developed to estimate missing measurements but there is no standard for this processing. Many gap-fill methods produce similar results on an annual basis but differences among methods may compound over a long period of time. In this study we applied four different gap-fill methods (two non-linear methods, and two variants of lookup table methods) to fill a 15-year continuous eddy-covariance dataset with the goal of determining differences in annual NEP. The dataset used was from a long-term flux tower site at the Northern Old-Growth Black Spruce (*Picea mariana*) forest site near Thompson, Manitoba, Canada. Over the 15-year dataset, annual NEP from the different methods was 4, 22, 44 and 48 g C m^{-2} y^{-1}, with the highest two values being significantly different than the lowest value. Overall, the four methods suggest the forest was a small carbon sink of 29 ± 10 g C m^{-2}

y^{-1}. Most significant differences in annual NEP occurred in the winter months and nighttime measurements, while significant differences in partitioned components, gross ecosystem production and ecosystem respiration, occurred mostly during the peak growing season.

2.2 Introduction

Boreal forests play a key role in the carbon balance of the northern hemisphere and consist of nearly one-third of the global land surface (Brandt 2009). The boreal biome spans a vast range across North America and Northern Europe encompassing a variety of continental climate regimes (Bergeron et al. 2007). Canadian boreal forests account for 34% of North American land cover (FAO 2010) and are responsible for nearly a quarter of global boreal forests (Brandt 2009). Boreal forests are in a particularly vulnerable location, expected to undergo rapid changes in climate (Soja et al. 2007). The response to climate change is uncertain and may vary based on region, making their study of great interest. It has been suggested globally that the majority of old-growth forests, including boreal forests, are currently carbon sinks (Luyssaert et al. 2008). In agreement, between the years of 1990 to 2008, the net ecosystem production (NEP, positive values are forest carbon gain) of Canadian managed boreal forests was approximately 28 ± 16 Tg C y^{-1} and overall were designated as a weak carbon sink (Kurz et al. 2013). It is estimated boreal forests globally contain nearly half of the terrestrial carbon sink, mainly due to the peat-rich soils contained in this biome (Gorham 1991). Since so much of the global carbon stores are invested in these forests, it is important to understand how they may change.

The northern old-growth black-spruce forest (NOBS) site is one of the most notable study sites for boreal research. This site was established in 1994 as part of NASA's Boreal Ecosystem-Atmosphere Study (BOREAS) equipped with a flux tower at the northern limit of the boreal forest in the Canadian Shield (Trumbore and Harden 1997). The site was commissioned to monitor biosphere-atmosphere interactions of a typical old-growth boreal forest (Sellers et al. 1995). This site has proven useful in many types of studies including model creation (Horn and Schulz 2011, Fu et al. 2014), model validation (Turner et al. 2006, Heinsch et al. 2006, Bonan et al. 2011, Stoy et al. 2014, Hilton et al. 2014), carbon flux analysis within the forest (Rocha et al. 2006, Dunn et al. 2007), response to environmental changes (Mahecha et al. 2010, Schwalm et al. 2010), and chronosequence analysis (Goulden et al. 2006, Goulden et al. 2011, Coursolle et al. 2012). Carbon balance at this site seems to indicate either carbon neutrality or a slight carbon source and is driven by site water balance (Dunn et al. 2007). A number of other boreal forest study sites are active in North America. The Southern Old Black Spruce (SOBS) forest site, also part of NASA's BOREAS study, is located in the Prince Albert National Park in central Saskatchewan, Canada. This forest, on an annual basis, was a moderate carbon sink (Krishnan et al. 2008). Another, in Alaska, was measured as a slight carbon source, and it was suggested the forest was undergoing effects caused by climate change (Ueyama et al. 2014).

Fluxnet is a global community involved in sharing datasets from hundreds of flux towers around the world. These sites range from oceans, to arable lands to forests. Flux is measured using the eddy covariance method which measures trace gas exchange between the atmosphere and an ecosystem (Baldocchi et al. 2008). These flux datasets are extremely useful in ecosystem science allowing inference of NEP as well as flux components, ecosystem respiration (R) – carbon loss, and gross ecosystem production (GEP) – carbon gained. Flux towers can be in place for a single growing season, or for a number of years. The possible longevity of datasets provides an excellent opportunity to study the ecosystem response to a changing climate, disturbance and other environmental pressures. Flux datasets are subject to gaps due to unsuitable weather conditions or instrumental failure/malfunction (Dragoni et al. 2007). Additionally quality control measures are used to ensure reliable data. Such measures usually include exclusion below a wind turbulence (u*) threshold (Goulden et al. 1997, Papale et al. 2006), NEP limits, and at times directional exclusions. Gaps make data less dependable and less useable (Hui et al. 2004), so gap-fill methods were developed to estimate missing data, yielding a complete dataset. It is not uncommon for data gaps to range from 30% to 70% of a flux dataset annually (Falge et al. 2001, Moffat et al. 2007) and most of these gaps will occur at night due to low turbulence (Moffat et al. 2007). For this reason, accurate gap-filling is extremely useful and is a step closer to being able to model 100% of carbon exchange.

There is currently no standardization for gap-filling flux data. It has been suggested standardization would improve comparability and increase the ease of synthesis activities from different biomes, ecosystem types and even to improve comparison among similar sites (Falge et al. 2001, Barr et al. 2002, Papale et al. 2006, Moffat et al. 2007, Desai et al. 2008). The issues arise first, when there are multiple gap-fill methods to select from, and second, when gap-filling methods are modified to suit study and even site need. Past studies have attempted to determine the uncertainty imparted by the gap-fill method. A few have compared many methods among a number of sites (Falge et al. 2001, Moffat et al. 2007) while others have compared fewer methods within a site (Ooba et al. 2006, Dragomir et al. 2012). It has been found that in comparing NEP filled with three different gap-fill methods, annual sums ranged between -200 to +45 g C m^{-2} y^{-1} (Falge et al. 2001). In a larger comparison, Moffat et al. (2007) determined gap-fill methods varied only ± 25 g C m^{-2} y^{-1} in annual NEP estimates. Within-site comparisons of gap-fill methods did not find a large range of NEP estimates, however these studies were short and involved only managed ecosystems (Ooba et al. 2006, Dragomir et al. 2012). It is not known how gap-fill method uncertainty affects flux datasets over a long period of time.

In the current study, we fill one of the longest available flux datasets spanning 15 years with four different gap-fill methods. The dataset was measured at NOBS from 1994 to 2008. Our objective was to determine the possible variability imposed by gap-fill method on the estimated annual NEP of NOBS. We selected four gap-fill methods, the Fluxnet Canada Research Network (FCRN) method (Barr et al. 2004), the Harvard method (Dunn et al. 2007), the method developed by the Max Planck Institute (MPI) of Biogeochemistry (Reichstein et al. 2005), and finally a method using dataset means to fill gaps, the Mean Data method. We hypothesized that there are no differences in NOBS annual NEP among methods.

2.3 Methods

2.3.1 Site Description

The NOBS site is located in the Canadian Shield at the northern limit of the boreal forest (Trumbore and Harden 1997); located at 55.88° N, 98.48° W in central Manitoba, Canada, about 40 km from the nearest city of Thompson, MB. There is no road access to the tower site so it must reached by either foot, all-terrain vehicle, or snowmobile; the closest road is approximately 4 km away (Provincial Highway 391).

Discontinuous permafrost underlies the soils at NOBS which were deposited by glacial Lake Agassiz. The soils are mainly clay and silt sediments and are peat rich, containing deep organic layers. Most of the landscape is flat but slight topographical changes create uplands and veneer bogs. The last recorded fire in the area was over 160 years ago (Gower et al. 1997). The uplands are well-drained and vegetation consists mainly of spruce trees (*Picea mariana*) averaging 10 m tall, and feathermoss (*Pleurozium* and *Hylocomium*). At lower elevations, the wetter, more poorly drained veneer bogs consist primarily of 1-6 m spruce and tamarack (*Larix laricina*) along with *Sphagnum* spp. The upland understory consists mainly of wild rose (*Rosa* spp.), and the veneer bog's understory consists of bog birch (*Betula glandulosa* var *hallii*), blueberry (*Vaccinium* spp.) and willow (*Salix spp.*), while Labrador tea (*Ledum groenlandicum*) is common throughout. Average stem density was 5450 tree ha^{-1} with a basal area of 35.6 m^2 ha^{-1} in 1994 (Gower et al. 1997). The sapwood volume was 82.5 m^3 ha^{-1} with a leaf area index of 4.2 (Gower et al. 1997). Surrounding the tower, within a 500 m radius, 50% of vegetation was classified as poorly drained characterized by veneer bogs (both feathermoss and *Sphagnum*), 25% were the moderately drained upland forests, and the final 25% very poorly drained fens (both *Sphagnum* and brown moss) (Harden et al. 1997).

2.3.2 Flux Measurements

Flux data were downloaded from Fluxnet archived datasets (ftp://daac.ornl.gov/data/fluxnet/fluxnet_canada/data/MB-NOldBlackSpruce/). The eddy covariance technique was used to calculate CO_2 turbulent fluxes on a half-hourly basis. The measurements were taken from a 31-m-tall flux tower, triangular in shape, each side measuring 30 cm. All other equipment, including data recording equipment, were kept in a hut 20 m away from the tower in order to protect from extreme temperatures. Primary and back-up power sources were provided to the site by two diesel generators 300 m east of the tower. Additionally, an uninterruptible power source (APC Smart-UPS 2200) was on site to permit recuperation from a power failure without need for human intervention. Data were recovered on a weekly basis.

Signals for turbulent flux calculations were recorded at 4 Hz on the tower at 29 m. A sonic anemometer (SATI/3K, Applied Technologies Inc., Boulder, CO) was used to measure the 3D wind velocities and temperature (T). Mixing ratios of CO_2, and H_2O were also measured at 29 m at a rate of 20 L min^{-1} through a 50 m long, 0.64 cm diameter Teflon PFA tube. A 4 L min^{-1} subsample was directed to a CO_2/H_2O infrared gas analyzer (IRGA; Model 6262, LI-COR Inc., Lincoln, NE). The voltages from the IRGA were used to calculate the gain of CO_2 every 3 hours by a standard addition of 4% CO_2 at 40 and 80 mL min^{-1}. Every 3 hours, an air sample through a CO_2 scrubber (soda lime) and desiccant ($Mg(ClO_4)_2$) was taken in order to determine the zero of the IRGA for CO_2 and H_2O, respectively. In order to determine canopy storage of CO_2, measurements were taken at 0.3, 1.5, 4.6, 8.4, 12.9 and 28.8 m sequentially at 0.5 Hz (half-hourly measurements) by a separate IRGA (Model 6262, LI-COR Inc., Lincoln, NE). Profile measurements were calibrated every 3 hours using two CO_2 mixtures traceable to National Oceanic and Atmospheric Administration (NOAA)/Climate Monitoring and Diagnostic Laboratory (CMDL) standards.

Net ecosystem exchange (NEE) was calculated as the sum of the measured eddy (F_c) and the air column storage (S_c).

$$NEE = F_c + S_c \qquad \text{(Eq. 2.1)}$$

Where S_c was determined as:

$$S_c = \int_0^{z_{ec}} \frac{\rho_a}{M_a} \frac{dC}{dt} dz \qquad \text{(Eq. 2.2)}$$

z_{ec} is the height above the ground of the eddy covariance measurement on the tower, ρ_a is the density of air when dry, M_a is the molecular weight of dry air and C represents the CO_2 molar mixing ratio. Controls were put on NEE in order to ensure quality data points. First, nighttime and low wind turbulence conditions were excluded below a u* threshold of 0.2 m s^{-1} (Dunn et al. 2007). Measurements of NEE exceeding ± 30 µmol m^{-2} s^{-1} were also excluded. Flux measurements recorded from the area of the diesel generators would have been biased so were excluded when wind came from between 45° and 135° from the north.

2.3.3 Supporting Meteorological Measurements and Variables

A quantum sensor (Model LI-190, LI-COR Inc., Lincoln, NE) at 29 m was used to measure incoming and outgoing photosynthetically active radiation (PAR). Also, eight sensors below the canopy measured PAR incoming to the forest floor. In winter, defined as air T < 0°C, GEP was set to zero, so gaps in PAR during these periods were not important. Gaps in PAR were filled first with PAR recorded at nearby tower sites. These consisted of nearby fen (55.91° N, 98.42° W) and young jack pine (55.90° N, 98.29° W) sites for years 1994-1996, UCI1930 (55.91° N, 98.52° W), UCI1981 (55.86° N, 98.48° W), UCI1989 (55.92° N, 98.96° W), UCI1964 (55.91° N, 98.38° W) for years 1999 and 2000, and finally wet and dry sites of both UCI1964 (55.91° N, 98.38° W and 55.92° N, 98.39° W respectively) and UCI1930 (55.90° N, 98.52° W and 55.91° N, 98.38° W respectively) for 2006 and 2007. Any remaining PAR gaps were filled with mean PAR from the exact time of the full dataset of 15 years. This may not reflect exact cloud cover but will reflect correct sun angle and day/night.

Soil T was also measured using five sets of thermistors at depths of 5, 10, 20, 50 and 100 cm. Soil T was not used in our analysis since over time the calibration of the soil T sensors shifted making measurements unreliable. For this reason we selected air T for all gap-filling. All the above-mentioned variables were recorded at 5 Hz. Air T was measured at 30 m; gaps were filled with air T measurements at 10 m, then 2 m. If gaps still existed, missing T was filled with T data from Thompson Airport, which correlated very well with NOBS T (Dunn et al. 2007). After this, very few data gaps (42 gaps of 262,992 data points) in air T existed, so any remaining gaps were filled by linear interpolation.

2.3.4 Gap-Filling Methods

2.3.4.1 Fluxnet Canada Research Network (FCRN) method (Barr et al. 2004)

First, short gaps in NEE, consisting of 4 half-hour periods or less, were filled using simple linear interpolation. This was done before filling larger gaps, which were filled using a “flexible moving window” to determine a time-varying parameter. This presumably allowed the model to account for environmental variability such as leaf area index, soil water content, and air saturation deficit as it changed with time. The window consisted of 100 data points, and these data points were used to calculate the slope of the linear regression between R and air T (Eq. 2.3), or PAR and GEP (Eq. 2.5) using only measured data (no gap-filled data). The slope was designated as the time-varying parameter for that window and was appointed the mean time of the 100 data points. The window was moved forward 20 data points at a time to determine the time-varying parameter along the dataset. Filling began by setting R equal to NEE at times when GEP was known to be zero (nighttime measurements and during the cold season). The following empirical, logistic relationship fitted estimated R to measured R values for the entire year:

$$R = f(T,t) = \frac{r_w(t) r_1}{1+\exp[r_2(r_3 - T_a)]} \qquad \text{(Eq. 2.3)}$$

where $r_{1,2,3}$ were empirical constants and $r_w(t)$ was the time-varying parameter estimated with the moving window where the linear regression was forced through zero. R (f(T,t)) corresponded to daytime estimates of R as well as missing nighttime R.

To fill GEP, the equation below was used for daytime measurements where measured NEE was available:

$$GEP = R - NEE \qquad \text{(Eq. 2.4)}$$

At nighttime, GEP was set to zero (when PAR is < 10 µmol m^{-2} s^{-1}). The following empirical model was fit to GEP (when not equal to zero):

$$GEP = f(Q,t) = \frac{p_w(t)\alpha Q \cdot P_x}{\alpha Q + P_x} \qquad \text{(Eq. 2.5)}$$

where α was the quantum yield, Q was PAR (µmol m^{-2} s^{-1}), P_x was the photosynthetic capacity (GEP at light saturation); both α and P_x were constants. The variable $p_w(t)$ was allowed to vary with time and was estimated using the moving window approach. The linear regression was also forced through zero. Note that GEP was only calculated using Eq. 2.5 when there was no measurement of NEE for use in Eq. 2.4.

In cases where there was no measurement for NEE, it was then filled by the subtraction of the calculated R and GEP using the following relationship:

$$NEE = R - GEP \qquad \text{(Eq. 2.6)}$$

Any leftover gaps (totalling 142 of the whole dataset 262,992 data points) were filled with the whole dataset mean at the exact time of the missing measurement. All programming was done using MATLAB (Mathworks Inc., Natick, MA, USA).

2.3.4.2 Harvard Method (Dunn et al. 2007)

Contrary to our standard quality control (Section 2.3.2), the NEE limits used by Dunn et al. (2007) were set to exclude NEE if below -17 µmol m^{-2} s^{-1} rather than exceeding ± 30 µmol m^{-2} s^{-1} and did not use a specific upper limit. This led to a difference of only 7 of 262,992 data points filled. Dunn et al. (2007) also filled PAR differently; by using a look-up table approach rather than filling with nearby sites first as we did in this study. In winter periods, classified by Dunn et al. (2007) as air T < - 10°C, there was no photosynthesis (GEP); so all flux measurements were considered to represent R flux. In order to obtain continuous R measurements for the whole dataset, intervals of approximately 10 days were chosen containing 100 accepted nighttime flux data points. The data points were tested to determine if the linear relationship with air T was significant. If it was, R was calculated with a linear function (to fill any nighttime gaps and daytime R values). If the regression was not significant, R was filled using linear interpolation using valid nighttime measurements. Linear regression fits more robustly than exponential fit over a short time period (as opposed to long periods of time) (Dunn et al. 2007).

The start and end of growing season was manually selected by the researcher (Dunn et al. 2007) each year by examining fluxes. The GEP was set to 0 until 14 days before the manually selected start of the growing season and 14 days after the selected end of the growing season. The start of growing season day of year, for 1994 to 2006 respectively, selected by Dunn et al. (2007) were 128, 121, 120, 143, 99, 103, 111, 115, 132, 110, 135, 96, 99 while the end of growing season day of years were 295, 295, 291, 293, 293, 290, 298, 298, 285, 297, 287, 308, 295. To estimate GEP, data were divided into periods of approximately 150 good GEP estimates (daytime, $u^* > 0.2$ m s^{-1} and T > -10°C). The data were tested for whether a linear or quadratic relationship based on PAR was representative. If the quadratic coefficients were significant, a model using nonlinear least squares regression was used for daytime GEP estimates with the following equation:

$$GEP = \frac{A+B(Q)}{(C+Q)} \qquad \text{(Eq. 2.7)}$$

where A, B and C are regression coefficients. If the quadratic relationship was non-significant, linear relationships were used for the same periods. Gaps in NEE were then filled with Eq. 2.6.

For the period of 1994-2004, we used the exact dataset filled with the Harvard gap-fill method as reported by Dunn et al. (2007) (ameriflux.lbl.gov). We had two extra years of data added to the Harvard dataset (2005 and 2006) also filled by Dr. Allison Dunn using the same method as outlined above.

2.3.4.3 Max Planck Institute of Biogeochemistry (MPI) Method (Reichstein et al. 2005)

This method was used by accessing the online gap-fill tool (http://www.bgc-jena.mpg.de/~MDIwork/eddyproc/index.php). As an exception to the other gap-fill methods; PAR was only used to distinguish between night and day measurements and was used as incoming solar radiation approximated by dividing PAR (µmol m^{-2} s^{-1}) by 2, for units of W m^{-2}. Small gaps in NEE were first filled with linear interpolation (4 half-hour periods or less). Data were then classified into three categories. In the first category, all meteorological variables were available, the only missing variable was that which was to be estimated (i.e NEE was the only missing variable). For this category, the gap was filled using the data point average of a 7-day window, as long as meteorological conditions were similar. Similarity was defined by incoming solar radiation measurements within 50 W m^{-2}, air T within 2.5°C, and vapour pressure deficit (VPD) within 5.0 hPa. If conditions did not meet these criteria within the 7-day window, the number of days was increased to 14 days.

In the second category, the variable of interest was missing, along with air T and VPD but radiation data were available. Gaps were filled in the same manner as the first category (average of a 7-day window) but similar conditions were defined only by radiation (within 50 W m^{-2}). In this case, if the 7-day window was not sufficient, the number of days was not increased.

Data fell into the third category when radiation data were also missing; these data were considered "poor quality". Gaps were filled with the average value at the same time of day (within ±1 hour) (mean diurnal course). The window size began at ± 0.5 days. If not able to be filled, window size was increased and the fill procedure was repeated with larger window sizes until all gaps were filled.

Flux-partitioning was based only on the original data (no gap-filled data). Nighttime data were selected with the radiation threshold of < 20 W m^{-2} and was then confirmed with standard sun geometrical routines. These NEE measurements corresponded to R. To fill gaps, data were split into periods of 10 days. The minimum number of data points within this window was 6 data points. The air T must also have a range of > 5°C allowing for proper regressions between air T and R. For each period, the Lloyd-and-Taylor (1994) regression was applied to determine short-term T sensitivity.

$$R(T) = R_{ref} \cdot e^{E_0\left(\frac{1}{T_{ref}-T_0}-\frac{1}{T-T_0}\right)} \qquad \text{(Eq. 2.8)}$$

The regression parameter (T_0) was set to -46.02°C. The activation energy (E_0) determines the temperature sensitivity of the regression and was allowed to vary over time. The reference temperature (T_{ref}) was set to 10 °C (to mimic the original model).

Periods of time when the standard error of the E_0 estimates were less than 0, and where estimates were realistic were accepted. Then the three periods with the smallest standard error were chosen to best represent the short-term T response of R. These three E_0s were then averaged and this was the value designated for the entire dataset ($E_{0,avg}$). To finish, R at T_{ref} was determined from the nighttime data for consecutive periods of days (4 days in this case). E_0 was set to $E_{0,avg}$ again in Eq. 2.8. The calculated R was then set to the central time point of the 4 day period and R was then interpolated between estimated points. In other words, for every half-hour, E_0 and R_{ref} were available, and were used to estimate R as a function of the air T used to derive the parameters. At this point, NEE and R were available for all time points, and GEP was calculated using Eq. 2.4.

2.3.4.4 Mean Data Method

The mean data method is an older method of filling data, also similar to mean diurnal variation or lookup tables (Falge et al. 2001). We took the entire 15-year dataset, filled any gaps in NEE with the average value of the variable of interest at the same day of year and time based on the entire dataset. Since this method did not have its own unique method of partitioning GEP and R we did not consider it in flux-partitioning analysis.

From this point forward, we converted NEE into NEP for all gap-fill methods, with the assumption that there is no loss of soil dissolved carbon in order to assure the following:

$$NEP = -NEE \qquad \text{(Eq. 2.9)}$$

2.3.5 Statistical Analyses

Annual cumulative NEP, GEP and R were compared among gap-filling methods using the Kruskal-Wallis non-parametric one-way ANOVA because data did not pass a test for normality (all annual NEP p-values > 0.2). Separate years for each method were used as replicates. Statistical tests only compared gap-fill methods for common years. Data from 1994 were not used in annual NEP, GEP or R analyses as that year was not a complete year of measurements. Years 2007 and 2008 were not included because those years were not available for Harvard-filled data. For GEP and R, data from 1996 were also excluded in statistical analysis because flux-partitioning was not available for MPI. Along with the annual comparisons, we included a hypothetic carbon neutral forest. This consisted of a dataset of 12 years where annual NEP was 0 g C m^{-2} y^{-1} in order to determine how each gap-fill method would compare to a forest with 0 net flux. Post-hoc comparisons were done using the Wilcoxon rank sum test for non-parametric data. We also compared average monthly cumulative NEP, GEP and R among methods in the same manner using the same common years. Finally, Kruskal-Wallis non-parametric ANOVA's were used to compare summer to winter (winter < 0°C) and day to night (night < 10 μmol m^{-2} s^{-1}) data for NEP, GEP and R among methods. For all statistical tests, an α-value of 0.05 was selected. All statistical analysis was performed using MATLAB software (Mathworks Inc., Natick, MA, USA).

2.4 Results

2.4.1 Environmental Conditions

Mean annual air T for NOBS during the study period was -2.5°C (1994-2008), which was slightly warmer than the Thompson airport historical mean annual air T for

1970-2013 (-2.9°C) (Figure 2.1). Mean air T ranged from -5.5°C in 1972 to -0.13°C in 2010. During the study period, the range was smaller, ranging from -4.4°C in 2004 to -0.8°C in 2006. Total precipitation for NOBS during the study period was 518 mm and was slightly less than the Thompson airport historical value of 525 mm for 1970-2013. The range of total precipitation values for both the historical time period and the study period were the same, from 368 mm in 2003 to a maximum of 894 mm in 2005.

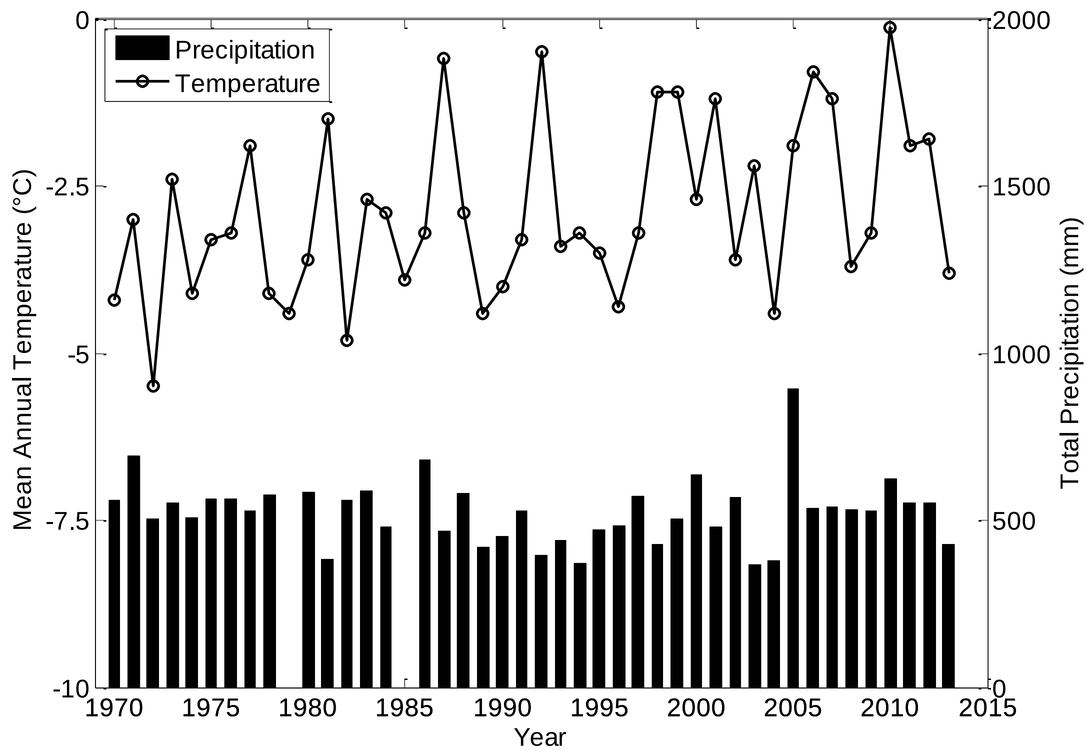

Figure 2.1: Historical air temperature and precipitation data (1970-2013) from Thompson Airport. Data points represent annual means for air temperature and bars represent annual total precipitation. Mean annual air temperature for 1970-2013 was -2.9°C while total annual precipitation was 525.3 mm. For the study period (1994-2008), mean annual air temperature was -2.5°C while total annual precipitation was 517.8 mm. Missing bars represent unavailable precipitation data.

2.4.2 Characterizing the Gaps

Datasets filled with different gap-fill methods were of the same size and underwent the same quality control measures, and had the same amount of gaps (Table 2.1). Before quality control, the raw dataset contained 31.7% gaps (26.1% from 1994-2006) while afterwards, the total percentage of gaps increased to 55.3% (52.1% from 1994-2006). The years 1994-2006 are mentioned since Harvard gap-fill data was only available until 2006. Most data gaps occurred during night-time measurements, and seasonally during winter measurements (Table 2.1). Among years, total gaps ranged from 47.6% in 2002 to 82.5% in 2007 (Figure 2.2). Year 2007 had an exceptional amount of gaps due to issues with the site generator. For winter measurements, gaps ranged from 51.2% in 2002 to 84.5% in 2007 whereas in the summer, data gaps ranged from 38.1 % in 1999 to 80.6% in 2007. Night-time measurement gaps ranged similarly to winter, 57.5% in 2002 to 85.6% in 2007 while daytime gaps ranged from 37.6% in 2002 to 79.5% in 2007 (Table 2.1).

Table 2.1: Total percentage of gaps from 14 year dataset (1995-2008) from the northern old-growth black spruce forest (NOBS) before quality control, and after quality control for wind direction, low turbulence, and net ecosystem production limits.

	Pre-Quality Control	**Post Quality Control**
Whole Dataset	31.7 %	55.3 %
Summer	26.4 %	52.6 %
Winter	40.6 %	65.1 %
Day	32.4 %	50.7 %
Night	35.5 %	67.8 %

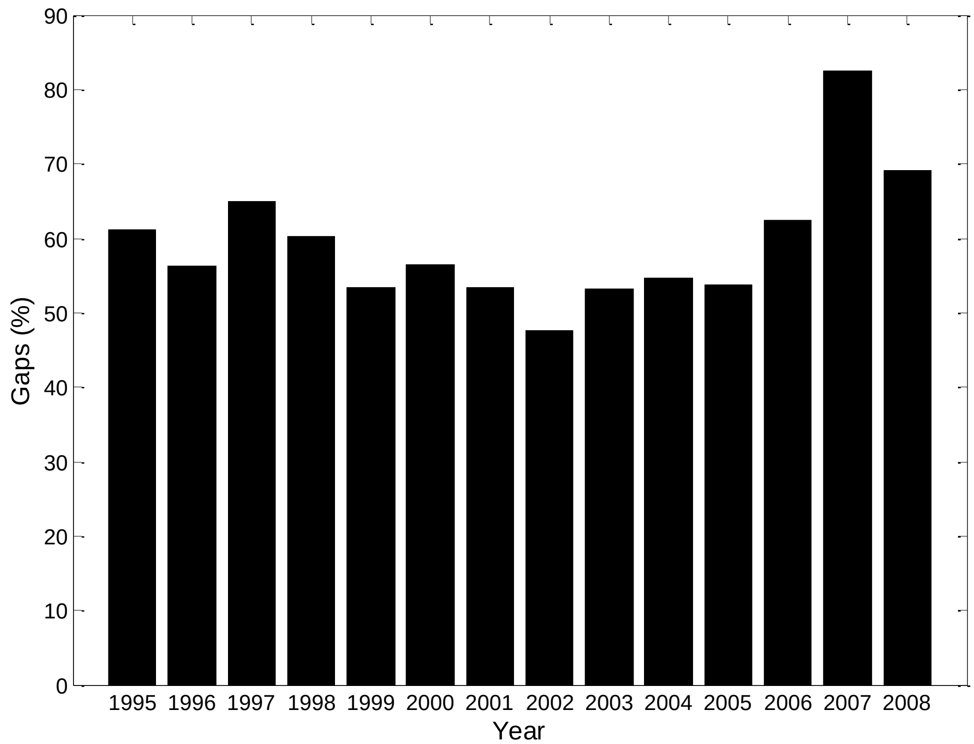

Figure 2.2: Percentage of gaps in dataset each year. Data were filtered for wind turbulence, net ecosystem production limits, and wind direction. The year 1994 was excluded in this figure since measurements did not begin until day of year 95.

2.4.3 Net Ecosystem Production

Although the actual magnitude of NEP was different among gap-fill methods, all seemed to follow a similar trend across the years of the study (Figure 2.3). From 1995 to 2003, in general, all methods estimated increasingly more positive annual NEP estimates. The only exception was in 2000 when FCRN estimated a lower NEP in relation to other methods. Harvard usually estimated the lowest annual NEP in any given year except for 2006 when it was among the largest with the Mean Data method. Typically, Harvard estimated only half the carbon uptake in comparison to the other three gap-fill methods. In 1995 and 1996 there were some opposing estimates among gap-fill methods: FCRN and Harvard estimated carbon losses, while MPI and Mean Data estimated carbon gains.

Harvard was the sole gap-fill method to estimate a carbon loss in 1997. FCRN was the only gap-fill method that estimated near-carbon neutrality/slight carbon loss in 2000.

Annual NEP was significantly different among gap-fill methods (χ^2 = 26.53, p < 0.0001) (Figure 2.3, Table 2.2). MPI and Mean Data gap-fill were significantly larger on an annual basis than Harvard. FCRN annual NEP was not significantly different from any other method. Annually, all methods tested as a significantly greater sink than the hypothetical carbon neutral forest (0 g C m^{-2} y^{-1}). Despite annual averages, values tested as significant sinks since in the majority of the years, cumulative annual NEP was above 0 g C m^{-2} y^{-1}. In the instance of Harvard and FCRN gap-fill, three quarters of the 12 years tested were above 0 g C m^{-2} y^{-1} cumulatively (p = 0.03); While in MPI and Mean Data method, all years were above 0 g C m^{-2} y^{-1} (p < 0.0001).

In winter, monthly NEP was always negative (October through to March) (Table 2.2). In April, monthly NEP became carbon neutral, followed by three months of peak monthly NEP. August brought a large drop in monthly NEP followed by a month of near carbon neutrality. In general, all methods agree on these trends. The months from January to August showed no significant difference in average monthly NEP. In October to December, FCRN-filled monthly NEP was significantly more positive than Harvard-filled data (χ^2 = 9.1, p = 0.03; χ^2 = 15.43, p = 0.002, χ^2 = 8.44, p = 0.04 respectively). In October, Mean Data was also significantly greater than Harvard; and in November, Mean Data was not only significantly more positive than Harvard, but also significantly more negative than FCRN. Mean Monthly Mean Data NEP was significantly greater than Harvard and FCRN monthly NEP in September (χ^2 = 9.88, p = 0.02). Most disagreement among gap-fill methods occurred at the end of growing season (September to December).

These results were reflected in comparisons among gap-fill methods when looking at all winter data together ($\chi^2 = 7.93$, $p = 0.047$; Table 2.3). It's important to note that although significant differences existed, the absolute fluxes during these periods were much smaller in magnitude than fluxes that occurred during the summer (e.g. May, June and July). When looking at all summer data, Mean Data annual NEP was significantly larger than both FCRN and Harvard gap-fill methods but no method was significantly different than MPI ($\chi^2 = 9.53$, $p = 0.023$; Table 2.3). When all daytime NEP data were compared among methods, there were no significant differences in estimates ($\chi^2 = 0.22$, $p = 0.97$; Table 2.4). Nighttime estimates produced significant differences, where FCRN and Harvard gap-fill estimated significantly more negative NEP than MPI and Mean Data ($\chi^2 = 17.13$, $p = 0.0007$; Table 2.4).

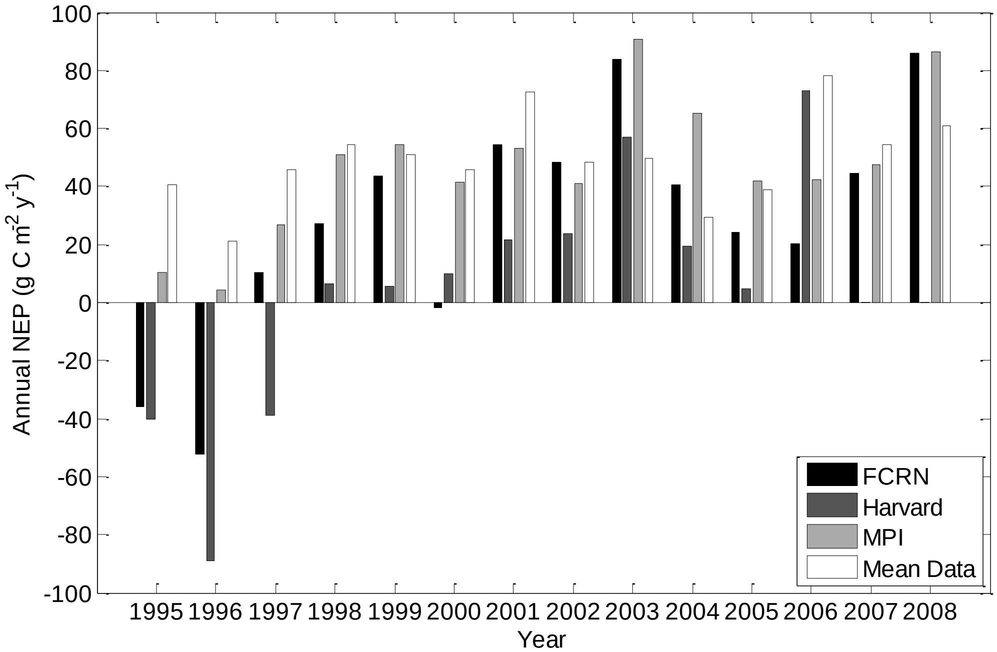

Figure 2.3: Annual net ecosystem production (NEP) from 14-year (1995-2008) eddy covariance dataset measured at the northern old-growth black spruce (NOBS) forest. Gaps in NEP were estimated using four different gap-fill methods before NEP was calculated: Harvard method, Max Planck Institute (MPI) method, Fluxnet Canada Research Network method (FCRN) and filling with Mean Data. The data from 1994 were excluded since measurements did not begin until day of year 95. No data were available for Harvard in 2007 and 2008.

Table 2.2: Average monthly net ecosystem production (NEP) from a 12-year (1995-2006) eddy covariance dataset measured at the northern old-growth black spruce (NOBS) forest. Gaps in NEP were estimated using four different gap-fill methods: Harvard method, Max Planck Institute (MPI) method, Fluxnet Canada Research Network method (FCRN) and Mean Data gap-fill method. Monthly NEP (within a single month/row) that are significantly different among gap-fill methods are indicated by letters based on a Wilcoxon rank sum test. Values are presented as means ± standard error (g C m^{-2} $month^{-1}$). Data from 1994 were not included since it was not a full year of data, and data from 2007 and 2008 were not used in the calculation of means because data were not available for all methods.

	FCRN		Harvard		MPI		Mean Data	
January	-7.2 ± 1.2		-7.2 ± 0.8		-6.5 ± 0.6		-6.7 ± 0.3	
February	-6.9 ± 1.9		-6.3 ± 1.0		-5.2 ± 0.6		-5.3 ± 0.2	
March	-5.8 ± 0.9		-9.3 ± 1.0		-7.3 ± 0.8		-7.4 ± 0.4	
April	0.1 ± 3.2		-4.7 ± 2.5		0.5 ± 3.8		-2.3 ± 2.3	
May	28.9 ± 4.2		26.8 ± 4.6		30.9 ± 4.2		32.3 ± 3.2	
June	33.6 ± 4.5		34.3 ± 3.9		38.5 ± 3.6		41.4 ± 1.5	
July	14.2 ± 2.7		13.5 ± 3.2		19.6 ± 1.9		20.8 ± 1.0	
August	-2.7 ± 4.2		-0.5 ± 4.7		6.6 ± 3.7		6.4 ± 2.2	
September	-1.2 ± 2.2	a	1.1 ± 2.1	a	2.5 ± 1.3	ab	5.9 ± 0.5	b
October	-11.2 ± 1.1	a	-16.1 ± 1.2	b	-12.1 ± 1.1	ab	-11.9 ± 0.7	a
November	-12.9 ± 0.8	a	-16.6 ± 0.7	b	-14.4 ± 0.5	ac	-14.9 ± 0.3	c
December	-7.6 ± 0.6	a	-10.9 ± 0.8	b	-9.6 ± 0.6	b	-9.8 ± 0.3	ab
Annual Mean (g C m^{-2} y^{-1})	21.9 ± 11.0	ab	4.5 ± 12.6	a	43.5 ± 6.7	b	48.0 ± 4.5	b

Table 2.3: Mean annual net ecosystem production (NEP) separated by summer and winter from 12 years of eddy covariance dataset measured at the northern old-growth black spruce (NOBS) forest. Flux data was filled using four different gap-fill methods: Fluxnet Canada Research Network (FCRN) method, Harvard method, Max Planck Institute (MPI) method and Mean Data method. The year 1994 is excluded for all gap-fill methods since there was not a full year of data. Data were not available for all methods in 2007 and 2008 so were excluded in order to eliminate bias. Means are expressed as ± SE (g C m^{-2} y^{-1}). Seasonal NEP (within a row) that were significantly different among gap-fill methods are indicated by letters based on a Wilcoxon rank sum test.

	FCRN		Harvard		MPI		Mean Data	
Summer	77.0 ± 10.7	a	71.9 ± 12.2	a	100.6 ± 6.3	ab	107.6 ± 4.4	b
Winter	-55.0 ± 3.8	a	-67.4 ± 3.6	b	-57.1 ± 2.8	ab	-59.6 ± 2.0	ab

Table 2.4: Mean annual net ecosystem production (NEP) separated by day and night from 12 years of eddy covariance dataset measured at the northern old-growth black spruce (NOBS) forest. Flux data was filled using four different gap-fill methods: Fluxnet Canada Research Network (FCRN) method, Harvard method, Max Planck Institute (MPI) method and Mean Data method. The year 1994 is excluded for all gap-fill methods since there was not a full year of data. Data were not available for all methods in 2007 and 2008 so were excluded in order to eliminate bias. Means are expressed as ± SE (g C m^{-2} y^{-1}). Day or night NEP (within a row) that were significantly different among gap-fill methods are indicated by letters based on a Wilcoxon rank sum test.

	FCRN		Harvard		MPI		Mean Data	
Day	254.3 ± 10.1		255.5 ± 9.2		249.5 ± 10.9		255.8 ± 6.6	
Night	-232.4 ± 9.7	a	-251.0 ± 7.0	a	-206.0 ± 10.0	b	-207.8 ± 4.4	b

2.4.4 Ecosystem Respiration

Partitioning of NEP into GEP and R helps define differences among methods and could also provide some clues to the differences. Mean Data method does not have an associated flux-partitioning routine for NEP so was not included in the analysis.

Annually, R was generally larger in the first part of the study period than the second half (Figure 2.4). All three methods estimated a decrease in annual R in 2004 followed by two years of larger annual R. The decrease experienced in 2004 was much greater when estimated by FCRN and MPI than by Harvard. MPI-partitioned R was the lowest annual R in all years except for 2007. Annually, R varied significantly among gap-fill methods ($\chi^2 = 14.08$, $p = 0.0009$) (Figure 2.4, Table 2.5). Both FCRN and Harvard were significantly greater than MPI method. In general there was less variability in annual R estimated by Harvard than by FCRN and MPI estimated R.

All methods agreed that R was low in the winter and peaked in the summer (July/August) (Table 2.5, Table 2.6). MPI-partitioned R was the lowest R in any given month except for November and December. In fact, MPI was significantly lower than Harvard in all months except for February, March and April ($\chi^2 = 9.04$, $p = 0.0109$; $\chi^2 = 8.68$, $p = 0.0131$; $\chi^2 = 13.21$, $p = 0.0014$; $\chi^2 = 16.8$, $p = 0.0002$; $\chi^2 = 10.45$, $p = 0.0054$; $\chi^2 = 6.18$, $p = 0.0455$; $\chi^2 = 10.85$, $p = 0.0044$; $\chi^2 = 9.96$, $p = 0.0069$; $\chi^2 = 8.27$, $p = 0.0166$ for January, and May to December respectively). MPI-partitioned R was significantly lower than FCRN-partitioned R in January, June, July and August. In October, November, and December, FCRN R was also significantly lower than Harvard. Most significant differences among gap-fill methods occurred during the summer months: MPI-partitioned R was significantly lower than FCRN and Harvard R ($\chi^2 = 13.28$, $p = 0.0013$) (Table 2.6). Over the winter months, Harvard had a higher R than the other two methods ($\chi^2 = 12.24$, $p = 0.0022$) but these fluxes were much smaller in magnitude compared to summer fluxes. MPI nighttime R was significantly lower than Harvard nighttime R ($\chi^2 = 12.6$, $p = 0.0018$) (Table 2.7). Daytime estimates also found MPI-partitioned R to be lower than both FCRN and Harvard-partitioned R ($\chi^2 = 14.3$, $p = 0.0008$) (Table 2.7). Overall, nighttime R was about 50 to 60% of daytime R. Harvard was significantly larger than both FCRN and MPI, while FCRN was significantly larger than MPI ($\chi^2 = 14.3$, $p = 0.0008$).

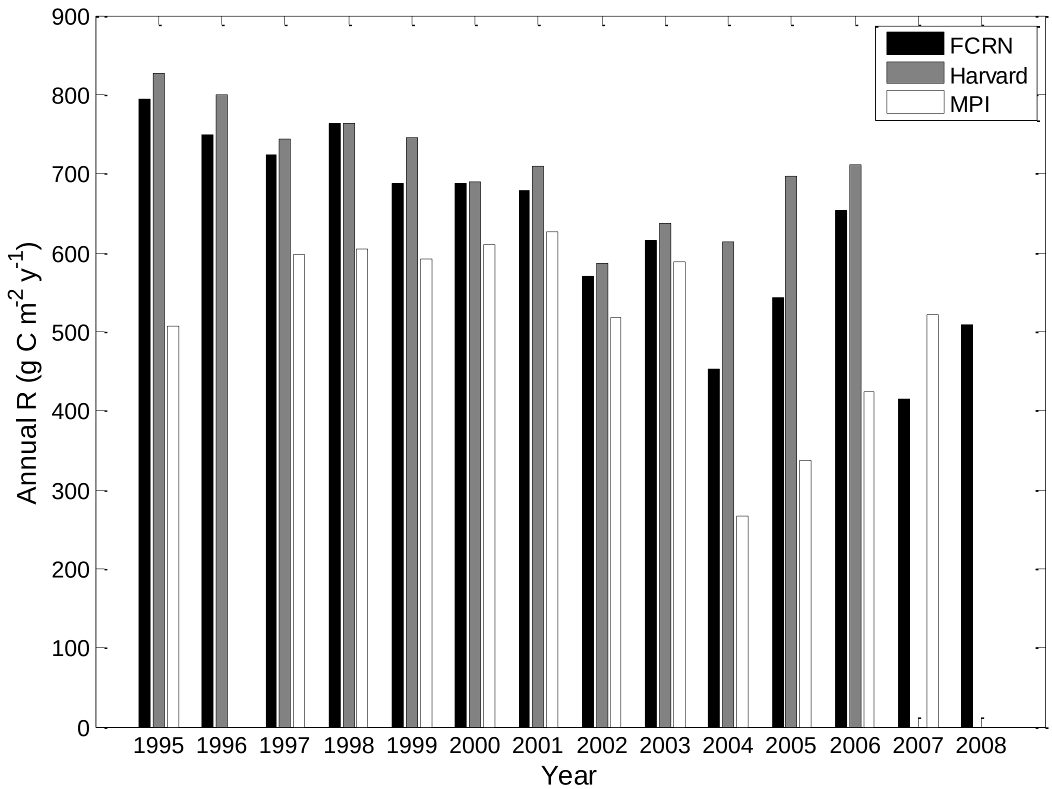

Figure 2.4: Annual ecosystem respiration (R) from a 14-year (1995-2008) eddy covariance dataset available for the northern old-growth black spruce (NOBS) forest. Estimates of R were determined using three different flux-partitioning tools: Fluxnet Canada Research Network (FCRN) method, Harvard gap-fill method, and Max Planck Institute (MPI) method. The data from 1994 were excluded since there was not a full year of measurements available. Flux-partitioning was not available in 1996 and 2008 for MPI method and not available in 2007 and 2008 for the Harvard Method.

Table 2.5: Average monthly ecosystem respiration (R) from a 12-year (1995-2006) eddy covariance dataset measured at the northern old-growth black spruce (NOBS) forest. Flux partitioning of R was performed using three different routines: Harvard method, Max Planck Institute (MPI) method and Fluxnet Canada Research Network method (FCRN). Monthly R (within a single month/row) that are significantly different among gap-fill methods are indicated by letters based on a Wilcoxon rank sum test. Values are presented as means ± standard error (g C m^{-2} $month^{-1}$). Data from 1994 were not included since it was not a full year of data, data from 1996 were excluded since flux portioning was not available for MPI method, and 2007 and 2008 were excluded in order to eliminate bias as data were not available for Harvard or MPI (2008).

	FCRN		Harvard		MPI	
January	6.9 ± 1.2	a	7.0 ± 0.8	a	4.6 ± 0.7	b
February	6.8 ± 2.1		6.2 ± 1.1		4.4 ± 0.8	
March	7.9 ± 1.2		9.6 ± 1.1		7.7 ± 1.5	
April	21.7 ± 2.3		22.9 ± 3.0		20.1 ± 2.7	
May	43.6 ± 3.5	ab	52.9 ± 4.7	a	34.7 ± 4.3	b
June	99.5 ± 6.7	a	108.2 ± 6.7	a	63.8 ± 9.2	b
July	153.1 ± 7.3	a	158.9 ± 5.7	a	111.8 ± 9.2	b
August	153.7 ± 8.6	a	157.4 ± 5.1	a	120.0 ± 9.7	b
September	91.6 ± 5.8	ab	95.9 ± 4.9	a	76.1 ± 5.6	b
October	32.4 ± 1.5	a	39.5 ± 1.3	b	31.9 ± 2.0	a
November	13.2 ± 0.9	a	16.7 ± 0.9	b	13.6 ± 0.7	a
December	7.3 ± 0.6	a	10.5 ± 0.8	b	9.9 ± 0.9	a
Annual Total (g C m^{-2} y^{-1})	660.0 ± 31.7	a	710.4 ± 22.2	a	515.7 ± 38.6	b

Table 2.6: Mean annual ecosystem respiration (R) separated by summer and winter from 12 years of eddy covariance dataset measured at the northern old-growth black spruce (NOBS) forest. Estimates of R were determined using three different flux-partitioning methods: Fluxnet Canada Research Network (FCRN) method, Harvard method, and Max Planck Institute (MPI) method. The year 1994 is excluded for all gap-fill methods since there was not a full year of data. Flux-partitioning was not available for MPI in 1996 and 2007 and 2008 were excluded in order to avoid bias. Means are expressed as ± SE (g C m^{-2} y^{-1}). Seasonal R (within a row) that are significantly different among gap-fill methods are indicated by letters based on a Wilcoxon rank sum test.

	FCRN		**Harvard**		**MPI**	
Summer	604.9 ± 28.4	a	633.8 ± 21.3	a	457.1 ± 33.8	b
Winter	55.0 ± 3.8	a	76.7 ± 4.1	b	58.6 ± 4.0	a

Table 2.7: Mean annual ecosystem respiration (R) separated by day and night from 12 years of eddy covariance dataset measured at the northern old-growth black spruce (NOBS) forest. Estimates of R were determined using three different flux-partitioning methods: Fluxnet Canada Research Network (FCRN) method, Harvard method, and Max Planck Institute (MPI) method. The year 1994 is excluded for all gap-fill methods since there was not a full year of data. Flux-partitioning was not available for MPI in 1996 and 2007 and 2008 were excluded in order to avoid bias. Means are expressed as ± SE (g C m^{-2} y^{-1}). Day and Night R (within a row) that are significantly different among gap-fill methods are indicated by letters based on a Wilcoxon rank sum test.

	FCRN		**Harvard**		**MPI**	
Day	427.1 ± 19.1	a	449.6 ± 13.9	a	319.2 ± 24.4	b
Night	232.9 ± 10.0	a	260.8 ± 7.4	b	196.5 ± 12.9	c

2.4.5 Gross Ecosystem Production

In general, both FCRN and Harvard flux-partitioning methods estimated larger GEP in the first half of the study (Figure 2.5). MPI annual GEP was largest closer to the mid-stage of the study period. All three flux-partitioning routines follow the same pattern from 2004-2006, increasing from lower to higher annual GEP throughout. The drop in 2004 was largest in MPI-filled and FCRN-filled data. Harvard only dropped slightly in comparison. MPI-filled annual GEP was the lowest GEP in any given year with the exception of 2007. On an annual basis, GEP varied significantly among gap-fill methods ($\chi^2 = 11.45$, $p = 0.0033$) (Table 2.8). All post-hoc tests agreed that Harvard and FCRN-filled annual GEP was significantly larger than MPI-filled annual NEP.

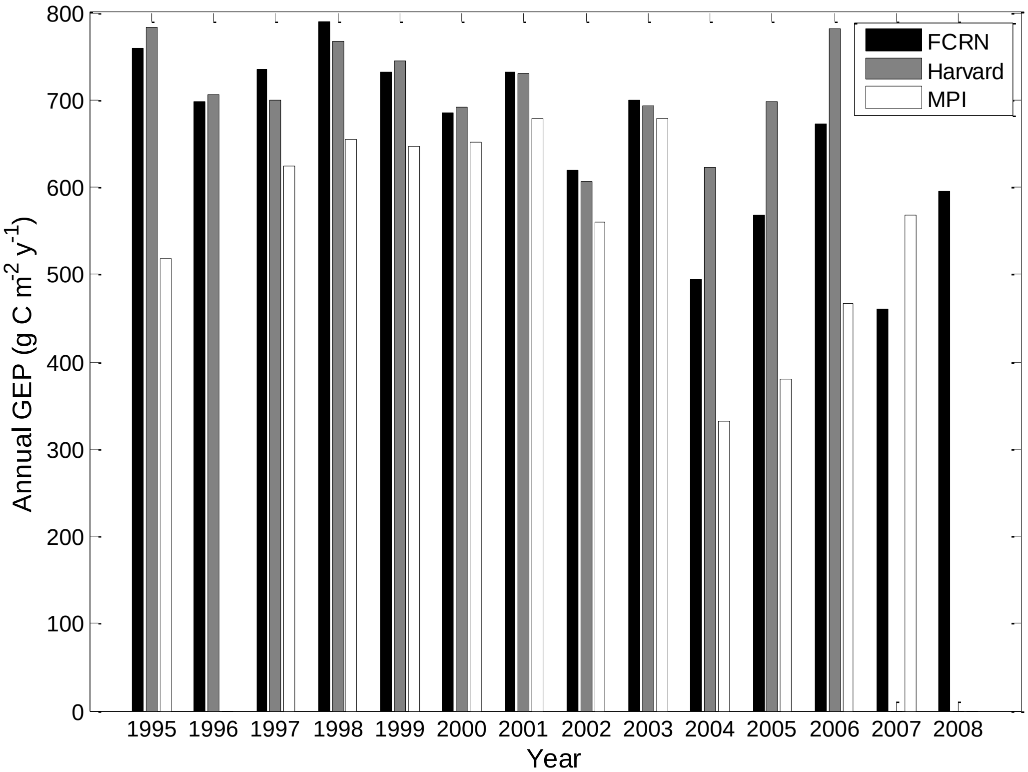

Figure 2.5: Annual gross ecosystem production (GEP) from a 14-year (1995-2008) eddy covariance dataset measured at the northern old-growth black spruce (NOBS) forest. Estimates for GEP were determined using three different flux-partitioning methods: Harvard method, Max Planck Institute of Biogeochemistry (MPI) method and Fluxnet Canada Research Network method (FCRN). The data from 1994 were excluded as it was not a full year of data. Flux-partitioning was not available in 1996 and 2008 for MPI method and not available in 2007 and 2008 for the Harvard method.

Table 2.8: Average monthly gross ecosystem production (GEP) from a 12-year (1995-2006) eddy covariance dataset measured at the northern old-growth black spruce (NOBS) forest. Flux partitioning of GEP was performed using three different routines: Harvard method, Max Planck Institute (MPI) method and Fluxnet Canada Research Network method (FCRN). Monthly GEP (within a single month/row) that are significantly different among gap-fill methods are indicated by same letters based on a Wilcoxon rank sum test. Values are presented as means ± SE (g C m^{-2} $month^{-1}$). Data from 1994 were not included since it was not a full year of data, data from 1996 were excluded since flux portioning was not available for MPI method, and 2007 and 2008 were excluded in order to eliminate bias as data were not available for Harvard or MPI (2008).

	FCRN		**Harvard**		**MPI**	
January	0.0 ± 0.0	a	0.0 ± 0.0	a	-2.0 ± 0.8	b
February	0.1 ± 0.1	a	0.0 ± 0.0	b	-0.9 ± 0.5	c
March	2.3 ± 0.7	a	0.1 ± 0.1	b	0.4 ± 1.1	b
April	22.6 ± 4.6		17.5 ± 5.1		22.7 ± 5.3	
May	74.1 ± 6.4		80.8 ± 7.9		67.6 ± 9.1	
June	136.7 ± 4.5	a	146.5 ± 4.0	a	104.5 ± 12.9	b
July	173.0 ± 6.2	a	178.0 ± 3.4	a	136.4 ± 13.8	b
August	156.6 ± 5.7	a	162.4 ± 1.9	a	132.7 ± 12.8	b
September	93.8 ± 5.1	ab	100.4 ± 3.9	a	81.3 ± 8.5	b
October	22.4 ± 1.8		24.3 ± 1.9		21.8 ± 2.5	
November	0.8 ± 0.2	a	0.4 ± 0.2	ab	-0.1 ± 0.5	b
December	0.0 ± 0.0	a	0.0 ± 0.0	a	-1.1 ± 0.8	b
Annual Total (g C m^{-2} y^{-1})	681.9 ± 27.9	a	710.2 ± 18.6	a	562.8 ± 39.1	b

Measurements of GEP should be zero in winter periods (i.e. November through to March for most years) (Table 2.8). Some methods estimated very small fluxes during these periods which were probably a result of either short periods when air temperature rose above freezing or noise in the NEE measurement. All methods agree upon GEP increases beginning in April, peaking in July and decreasing until November. Monthly GEP estimated by the MPI method was usually the lowest, with a few exceptions. There were no significant differences in monthly GEP among gap-fill methods in April, May and October. Most significant differences among monthly GEP occurred during the peak growing season (June through September). MPI-partitioned monthly GEP was significantly lower than Harvard-filled in January, February, June, July, August, September and December (χ^2 = 18.41, p = 0.0001; χ^2 = 3.95, p = 0.01; χ^2 = 11.82, p = 0.0027; χ^2 = 14.79, p = 0.0006; χ^2 = 15.07, p = 0.0005, χ^2 = 6.78, p = 0.0337; χ^2 = 10.15, p = 0.0062 respectively) (Table 2.8). Additionally, MPI-partitioned annual GEP was significantly lower than FCRN-partitioned GEP in January, February, March, July, August, September, November and December. FCRN was also significantly larger than Harvard in February. In the instance of GEP, we did not compare seasonal and diurnal GEP as with R; this is due to the zero values of GEP in both winter and at night. This means daytime and summer GEP are reflected, almost exactly, by the values we have evaluated here.

2.4.6 Cumulative NEP

Over the entire dataset, there was a large divergence in cumulative NEP among gap-fill methods (Figure 2.6). Both FCRN and Harvard estimated carbon loss in the early years of the study. This decreased to a low value of approximately -175 g C m^{-2} in 1996/1997 for FCRN and to a low value of approximately -300 g C m^{-2} for Harvard after which carbon began to accumulate in both instances. After 13 years of measurements (up to 2006), FCRN estimated NOBS carbon balance to be approximately 225 g C m^{-2} and Harvard estimated a slight source, -33 g C m^{-2}. After 15 years, FCRN estimated a final carbon balance of 373 g C m^{-2}. Using both MPI and Mean Data gap-fill methods, NEP climbed steadily from beginning to end of the study period. These estimates were similar and a final carbon balance of 695 g C m^{-2} and 768 g C m^{-2} were reached over 15 years for MPI and Mean Data methods, respectively. Over 13 years (up to 2006), these methods estimated carbon balance of approximately 550 and 600 g C m^{-2} respectively.

FCRN and Harvard followed the same pattern, but diverged around year 1996/1997. The difference early in the dataset was approximately 100 g C m^{-2} at any given time point. This difference appeared to increase as time went on to approximately 200 g C m^{-2} near year 2000 and reached nearly 300 g C m^{-2} in 2006 after 13 years. The largest differences among gap-fill methods were between Harvard and Mean Data totalling nearly 750 g C m^{-2} in 2006. Similarly, the disagreement between Harvard and MPI was about 700 g C m^{-2} in 2006. FCRN and Harvard differed by approximately 200-275 g C m^{-2} in 2006. By 2008, Mean Data and MPI were very similar and estimated approximately 400 g C m^{-2} more NEP than FCRN.

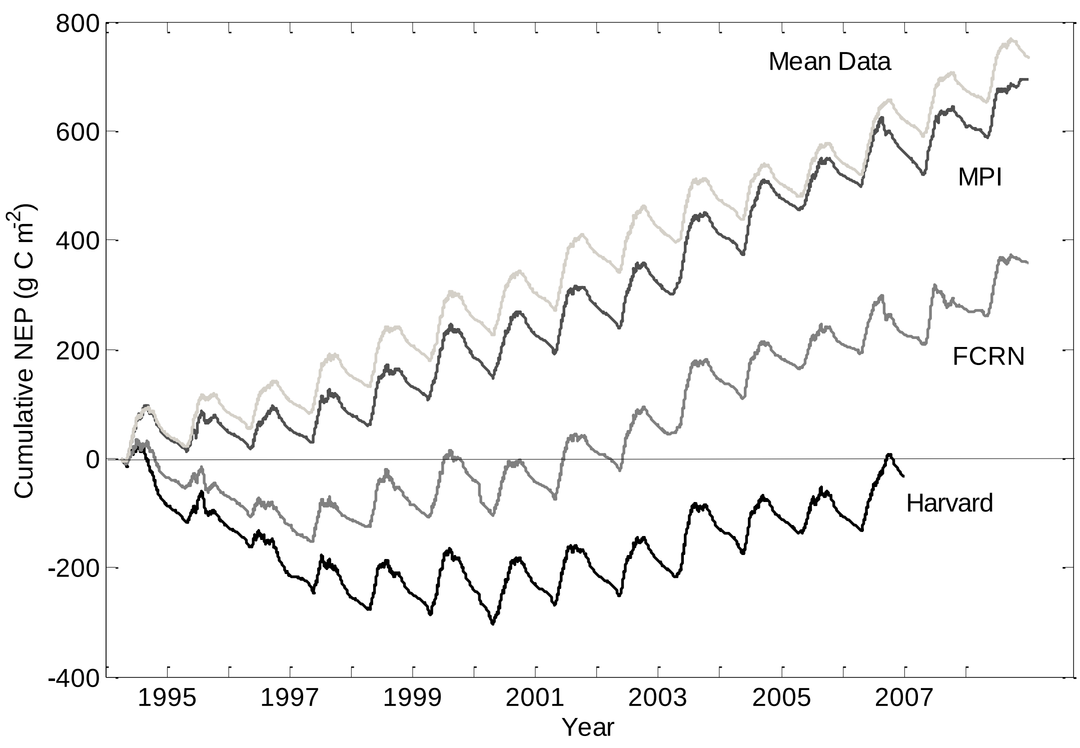

Figure 2.6: Cumulative net ecosystem production (NEP) from 15-year eddy covariance dataset available for the northern old-growth black spruce (NOBS) forest. Gaps in NEP were estimated using four different gap-fill methods: Harvard method, Max Planck Institute (MPI) method, Fluxnet Canada Research Network method (FCRN) and filling with Mean Data. Measurements did not begin until day of year 95 (in 1994) and data were not available for 2007 and 2008 for the Harvard gap-fill method.

2.5 Discussion

2.5.1 Differences in Annual NEP Among Gap-Fill Methods

2.5.1.1 Annual NEP

Annually, there was a clear difference in average NEP among gap-fill methods for 1995 to 2006. Both MPI and Mean Data gap-fill estimated on average the strongest sink (43.5 ± 6.7 g C m^{-2} y^{-1} and 48.0 ± 4.5 g C m^{-2} y^{-1} respectively). The lowest average annual NEP was estimated using the Harvard gap-fill method (4.5 ± 12.6 g C m^{-2} y^{-1}) while FCRN estimated in the middle (21.9 ± 11.0 g C m^{-2} y^{-1}). In comparison to other studies on NOBS, values agree depending on which gap-fill method was used. For example Dunn et al. (2007) used the Harvard gap-fill method and obtained an annual NEP average of 1.8 g C m^{-2} y^{-1} (1995-2004). Since the dataset in the current study was the same as used in the Dunn et al. (2007) paper, with an additional two years of data (2005 and 2006), this is not surprising. Using the FCRN gap-fill routine, Bergeron et al. (2007) estimated an annual NEP of 27.0 ± 11.0 g C m^{-2} y^{-1} in 2004, similar to our average annual NEP over 12 years (1995-2006) filling data with FCRN (in comparison to our value 21.9 ± 11.0 g C m^{-2} y^{-1}); However, for 2004 specifically in the present study, annual NEP was slightly lower (19.5 g C m^{-2} y^{-1}), but still fell within error estimates. Both MPI and Mean Data gap-fill estimated substantially higher than both of these studies annually.

The annual carbon balance of NOBS was in the range of 4 to 48 g C m^{-2} y^{-1}. As a mean of all gap-fill methods, we estimate NOBS to be a slight carbon sink of 29 ± 10 g C m^{-2} y^{-1}. Krishnan et al. (2008), measured an average annual NEP of 56.0 ± 21.0 g C m^{-2}

y^{-1} over 7 years at SOBS forest; but the site was warmer annually (0.4 °C) and received less precipitation (467 mm y^{-1}) than NOBS. Although the range of possible values was large for NOBS annual NEP, all estimates were lower than that of SOBS. In reality, the only reliable comparison that could be made between these two sites was FCRN gap-filled NOBS data since FCRN gap-fill was also used by Krishnan et al. (2008). This means both datasets will have similar errors and biases imposed on the data by gap-fill method, although many other forms of error exist (Desai et al. 2008). In an Alaskan old-growth black-spruce forest with similar mean annual temperature (-2.9 ± 1.3 °C) and much lower precipitation (263 ± 60 mm y^{-1}), over 9 years annual NEP averaged -11 g C m^{-2} y^{-1} (Ueyama et al. 2014); a slight carbon source. In comparison to NOBS average annual NEP of 29 g C m^{-2} y^{-1} (a slight sink), the forests are in differing states of carbon balance. It should be noted that a different gap-fill method was used by Ueyama et al. (2014) that was not compared in our study.

2.5.1.2 Effects on Cumulative Carbon Balance

Our study suggests that using these four gap-fill methods yield three different carbon balances for NOBS over a 15-year dataset. Since the pattern of FCRN-estimated cumulative NEP is similar to that of Harvard gap-fill, this indicates FCRN and Harvard gap-fill may respond to driving variables similarly. Dunn et al. (2007) suggested that water table depth and moisture regime control R of NOBS which in turn drives NEP. They found that climatic variables could not explain all the variation in NEP. Dunn et al. (2007) describe R as large in earlier years following a pre-study drought which caused the carbon loss at the beginning of the study. FCRN-partitioned R is very similar to Harvard R throughout the study until 2004. Years 2003 and 2004 were dry years at

NOBS and Dunn et al. (2007) reported no change in R, however, in both FCRN and MPI data, R decreased dramatically then increased in following years. Similar responses were seen in site GEP among all gap-fill methods. Based on NEP, we know all methods predicted carbon gains regardless of changes in GEP and R. Year 2005, following two dry and cold years, had among the lowest NEP values through the second half of the study for all methods excluding Mean Data gap-fill. It was suggested that the shift from a drought to above-average levels of precipitation and T is what caused NEP to finally begin to climb using the Harvard gap-fill method (Dunn et al. 2007). Long-term changes in the water table were thought to be responsible for the dip in NEP at the beginning of the study since one or two years of low rainfall (2003 and 2004) did not yield the same magnitude in change of NEP (Dunn et al. 2007). The differences between FCRN and Harvard cumulative NEP seem to stem from slight differences in how variables were derived. It is possible that Harvard predicts greater carbon losses if it is more sensitive to conditions causing R to be large as suggested by Dunn et al. (2007), in this case, to drought years.

This contrasts both MPI and Mean Data gap-fill methods in the overall study carbon balance. These two methods demonstrated continual accumulation throughout the 15-year study period, and never dipped into negative overall NEP. In general, Mean Data estimates slightly higher NEP throughout the study than MPI. MPI does show lower NEP in the 1995 to 1997 period when looking at NEP for each year (Figure 2.3), but remains more positive than Harvard or FCRN. This does not translate to the same pattern observed by FCRN and Harvard gap-fill for cumulative carbon balance, but it does seem MPI picks up on some of the same drivers as the other two gap-fill methods. The 15-year cumulative carbon balance estimated by MPI and Mean Data would not support changes observed in carbon balance associated with water balance observed by Dunn et al. (2007) as strongly. This lends more support for the prediction that rising temperature will/has increased ecosystem productivity (Briffa et al. 2008) since there was a slight trend in increasing temperatures in comparison to the 30-year temperature mean (Dunn et al. 2007). MPI and Mean Data estimate very similarly since both filled gaps with some sort of mean value. Mean data gap-fill takes the full dataset mean for the exact time of the gap, while MPI gap-fill calculates the mean of a 7-day window, averaging values from surrounding days. The Mean Data method estimates a larger NEP, as it will be biased towards the overall mean and would not capture environmental time series as well.

It has been previously suggested that short-term flux measurements and observations may not always be representative of long-term variations in flux (Ueyama et al. 2014). For example, in previous NOBS studies based on three years of data, increases in R were attributed mainly to growing season warming (Goulden et al. 1998). Goulden et al. (1998) suggested the sensitivity of R was due to an increase in the soil active layer (less frozen soil). Site R did not vary with precipitation or drying of soil during the short study period (Goulden et al. 1998). Dunn et al. (2007) studied the same site over a longer period of time (1994-2004) and were able to attribute the increases in R to associated changes in NOBS water balance. The water table changes act on longer time scales than could be observed by Goulden et al. (1998) so a different observation was made. This demonstrates the importance of long-term flux datasets such as is available for NOBS so that long-term climatic and environmental trends can be better attributed to measured fluxes and better predictions for future changes. That being said, it appears that choice of data processing may also affect interpretation of fluxes. This has previously been suggested in the importance of standardizing post-processing routines in order to increase abilities to compare among sites (Falge et al. 2001, Barr et al. 2002, Moffat et al. 2007). The other option here is that in customizing gap-fill methods, we may get more accurate estimates of carbon exchange. In this case, reporting uncertainty would be pertinent.

2.5.2 Temporal Differences in NEP

2.5.2.1 Time of Year

The monthly NEP of NOBS followed a pattern typical for northern boreal forests as well as other similar boreal forest types with the highest net carbon uptake in May and June (Goulden et al. 1998, Falge et al. 2002, Griffis et al. 2003, Bergeron et al. 2007, Dunn et al. 2007). This is due to a period of time when air warms much faster than soils, and causes a decoupling of R and GEP components (Falge et al. 2002, Bergeron et al. 2007, Dunn et al. 2007). So R is still low when soil T is cool but PAR is larger and drives changes in GEP. The carbon balance of NOBS also showed the characteristic mid-summer depression in NEP seen in boreal coniferous stands due to increases in R beginning in July (Griffis et al. 2003, Black et al. 2005, Bergeron et al. 2007). This occurs due to a reduction in GEP with shorter photoperiod (Suni et al. 2003). Partitioning of R and GEP at NOBS yielded characteristic annual “bell” curves caused by both the thawing of soils in the spring, increasing heterotrophic R (Goulden et al. 1998); and T increases allowing assimilation to begin with longer photoperiod (Suni et al. 2003). This was similar to trends seen in NOBS, SOBS, the eastern old-growth black spruce forest (EOBS) as well as a jack pine forest nearby in Saskatchewan (Goulden et al. 1998, Griffis et al. 2003, Bergeron et al. 2007, Dunn et al. 2007).

Deviating from the results in the present study, all three boreal forests in Bergeron et al. (2007) appear to have a large drop in NEP come July and a new increase in August, followed by a drop in NEP to near neutral in September. Data were gap-filled with FCRN and only used one year of data (2004). In 2004, using any of our gap-fill methods, there was a drop in NEP estimated by all gap-fill methods in July, but neither Harvard nor Mean Data estimated an increase following in August. FCRN and MPI both estimated an increase in NEP come August, but on a scale of 1 or 2 g C m^{-2} y^{-1} as opposed to the 10-15 g C m^{-2} y^{-1} increase suggested Bergeron et al. (2007). This could be an anomalous year in comparison to our 12-year averages that display a smoother decline in NEP from July to September in all four gap-fill methods. This demonstrates the importance of long-term studies in order to obtain a proper sample of all types of years (2004 was a dry year). Similarly, Goulden et al. (1998) and Dunn et al. (2007) found NOBS to be approximately carbon neutral beginning in July, and remained this way until winter when R increased. During the growing season, carbon accumulation is mediated by site water balance, water table and summer radiation (Baldocchi et al. 2008). In general, during growing seasons, gap-fill methods seem to respond similarly to these forcing variables. In August and September (end of growing season), FCRN and Harvard gap-fill estimated negative NEP on average, while MPI and Mean Data remained with positive NEP at this time although values were small.

Most significant differences in NEP among gap-fill methods occurred at the end of growing season (September and October). That being said, early and late growing season fluxes are small in comparison to fluxes measured at peak growing season and significant differences may only represent a few grams of carbon over the entire year.

Although Dunn et al. (2007) determined growing seasons manually (Harvard gap-fill data), this did not create a substantial difference in total NEP when compared to other gap-fill methods if imposed with the same growing season length and did not explain differences over time. Similarly, when looking at differences in total annual GEP, NEP and R among three boreal forest systems across different areas of Canada, Bergeron et al. (2007) determined that between two different methods of determining growing season, there was no difference in carbon balance observed within any of the three forests, including NOBS, SOBS and EOBS forest.

During winter, boreal forests will lose carbon very slowly, but this does contribute a significant amount with regards to the carbon balance of the site (Baldocchi et al. 2008). Although significant differences were seen when all summer data were compared (Table 2.3, Table 2.6), this was not observed in monthly comparisons of summer months. Many differences among gap-fill methods seem to be apparent during the winter when taking into account all data (Table 2.3, Table 2.6) and when looking at monthly comparisons (Table 2.2). During these times GEP is negligible, so NEP is driven by R. In the present study, Harvard gap-fill predicts the largest loss of carbon in the winter (highest R values) on average, approximately 20 g C m^{-2} y^{-1} greater loss than FCRN and MPI. Over 15 years, this could total upwards of approximately 300 g C m^{-2} y^{-1} and may explain some of the divergence between Harvard and FCRN since day, night and summer R estimates are similar between the two. In winter NEP, these differences were approximately 10 g C m^{-2} y^{-1}, but the only significant result was Harvard as a significantly greater loss than FCRN. This observation differs from what was observed in agricultural sites (Dragomir et al. 2012). In comparing three gap-fill methods, it was observed that there was greater accuracy among gap-fill methods in winter NEP estimates (Dragomir et al. 2012). It should be noted that winter NEP is much smaller (near zero/slightly negative) when compared to summer NEP which may have contributed to this observation.

2.5.2.2 Day Versus Night

Differences occurred among gap-fill methods mostly during daytime estimates in comparison to nighttime estimates in a study involving multiple forested sites and a number of gap-fill methods (Moffat et al. 2007). Similarly, errors in gap-filling were larger in daytime compared to nighttime (± 0.83 g C m^{-2} and ± 0.52 g C m^{-2} per % of year filled respectively) in a study by Falge et al. (2001) encompassing many sites from FLUXNET and a number of gap-fill methods. Daytime estimates differed significantly in this study as well but only with regards to R; both Harvard and FCRN were significantly higher than MPI by approximately 130 and 115 g C m^{-2} y^{-1} respectively (Table 2.7). There were no significant differences in daytime NEP. There were significant differences in nighttime NEP, where differences were as large as 45 g C m^{-2} y^{-1}. Incidentally, this difference from max to min was similar to that of average annual NEP for the full study. Nighttime NEP and R estimates may contribute to the differences observed annually among methods.

2.5.3 Flux-Partitioning

Flux partitioning is important in eddy covariance analysis, especially in long-term studies, in order to better understand interannual variability in site NEP (Valentini et al. 2000). It is required in understanding plant and soil responses and processes individually rather than whole ecosystem changes as there is not a direct measurement for either GEP or R. It is also important in diagnostics to determine which component may be misrepresented (Falge et al. 2002, Reichstein et al. 2002); because if both components are overestimated or underestimated, it may not be translated into a difference in NEP. Most commonly, GEP and R estimates are determined by extrapolating nighttime R measurements (= nighttime NEP) into daytime periods. There are different techniques within this broader category and it is possible that the slight modifications may be responsible for some of the differences seen among gap-fill methods, since each will partition NEP components differently. We hypothesized that analysis of these components may direct attention to where some estimates differ. In theory, since all three flux-partitioning routines attempt to account for environmental changes by using a variable that varies over time, similar results should be seen across all methods as they all used the same input variables (PAR and T air).

2.5.3.1 Ecosystem Respiration

As opposed to NEP, most significant differences in R occurred during growing season rather than during the colder season when fluxes were small. Significant differences almost exclusively distinguished MPI-partitioned R as lower than either Harvard and/or FCRN R while differences between Harvard and FCRN were only significant in certain cold months (Harvard was greater than FCRN). In a study comparing 23 flux partitioning methods on 10 years of data from multiple datasets, it was determined that 75% of methods fell within 100 g C m^{-2} y^{-1} for respiration (Desai et al. 2008). This 100 g C m^{-2} y^{-1} range is true between FCRN (660 ± 2 g C m^{-2} y^{-1}) and Harvard (710 ± 2 g C m^{-2} y^{-1}) but not if MPI is taken into account (517 ± 2 g C m^{-2} y^{-1}). Years 2004 and 2005 were the only times FCRN and Harvard deviated by more than 100 g C m^{-2} y^{-1}. As seen in Desai et al. (2008), methods in the current study also agreed well on seasonal trends with regards to both GEP and R. Diurnally, GEP usually agrees well, while R was more variable (Desai et al. 2008). We did not see this translate into monthly, annual or seasonal differences among gap-fill methods in the current study.

Studies directly analyzing flux-partitioning methods within and among sites are scant (Stoy et al. 2006, Desai et al. 2008) and are not unanimous in conclusions. For short-term R estimates, all methods performed poorly, however, complex partitioning methods yielded best results for long-term R variability (Stoy et al. 2006). The more complex methods used daytime flux data in an attempt to estimate R with short time windows (Stoy et al. 2006). In contrast, Desai et al. (2008) received better coherence with synthetic data from ecosystem carbon models by extrapolating nighttime measurements of R using a short-term T sensitivity, similar to that used in the MPI gap-fill method (Reichstein et al. 2005). In the present study, all three flux-partitioning routines extrapolated R from nighttime estimates, which are at least supported by one study as superior in comparison to synthetic data (Desai et al. 2008). Using the short-term T sensitivity in MPI gap-fill does have the short-coming of not accounting for changes in seasonal T sensitivity (Reichstein et al. 2005). This may be better accounted for in the other methods using interpolation of a set amount of data points rather than a years' worth of measurements, although it may not represent long-term T sensitivity as well (Reichstein et al. 2005). Interpolation using the 100-point moving window in FCRN gap-fill or 10-day window in Harvard gap-fill should account for long-term temperature sensitivity, although FCRN may be able to better account for changes over time due to the "moving window" nature of the method.

MPI gap-fill does not have a way to fill gaps that cannot be determined with R interpolation. In theory, R and the E_0 should be available for every half-hour period in the dataset and can be used to determine NEP and GEP. We did encounter two years of data where the MPI flux-partitioning tool was not able to complete the data. This is likely due to a few periods of data noisiness during those years or the large range of temperatures at the site where temperature sensitivity could not be derived (Reichstein et al. 2005).

It is unknown which T is best suited for gap-filling (air T, soil T at different depths or surface T) in order to model best responses (Subke and Bahn 2010). In one study comparing T inputs, air T tended to give higher estimates of R with all flux-partitioning methods than soil T (Lasslop et al. 2012). This is probably due to the larger diurnal fluxes observed in air T in comparison to soil T. The lag between soil and air T was always able to account for differences in estimates filled using different air variables (Lasslop et al. 2012). It was recommended that site-specific analysis be done to determine whether soil or air T best accounted for variability in R (Lasslop et al. 2012). In our study, we did not have the option of gap-filling with soil T since sensors were unreliable over time. It is however important to note the possibility of overestimation in general.

2.5.3.2 Gross Ecosystem Production

At nighttime and in winter (frozen conditions), GEP is assumed to be zero. MPI does not seem to have a set filter as with FCRN and Harvard for when GEP should be set to zero (either a temperature or PAR limit) and tends to estimate negative GEP in winter months, but this is small numerically (Table 2.4). Again, as with R, MPI-partitioned GEP is almost always significantly lower than either FCRN or Harvard GEP throughout the growing season. This does not translate into similar differences in NEP for any method. FCRN and Harvard gap-fill routines will use PAR to estimate GEP when flux measurements are missing. The functions used by each are slightly different, and again this may cause some of the divergence between the magnitudes of fluxes. Much of GEP is calculated as the residual of NEP and R (when NEP measurements are available). In fact, this is the only way GEP is calculated by the MPI flux-partitioning routine. In these cases, GEP will be biased by what NEP and R were already determined to be and may not reflect PAR driven responses in GEP.

2.5.4 Gap-Filling Methods and Ecosystem Modelling

Gap-filling methods have a wide variety of approaches that involve one or many of interpolation, look-up tables, non-linear regression, neural networks or process-based models (Falge et al. 2001, Moffat et al. 2007). This range of techniques is further complicated with site-specific modifications including different u* thresholds, measurement filters, whether measurements are half-hourly or hourly and site characteristics. All these differences create uncertainty and bias in each dataset making it very hard for data synthesis and inter-site comparison (Morgenstern et al. 2004) and as we show here, even within-site comparisons. Choice of gap-fill method and further modification can bias estimates towards certain climatic variables or environmental factors which can compound further into flux components. This emphasizes the importance of standardized processing methods (Falge et al. 2001, Barr et al. 2002, Papale et al. 2006, Moffat et al. 2007, Desai et al. 2008).

In comparing three gap-fill methods used to fill a dataset from May to September at a Japanese Larch (*Larix kaempferi*) plantation it was found that daily NEP was similar between data filled with neural network methods (-3.48 g m^{-2} d^{-1}) and filled with empirical equations (-2.76 g m^{-2} d^{-1}) (Ooba et al. 2006). It was however concluded that genetic neural network and artificial neural network methods were of a higher standard (Ooba et al. 2006). This was true if we considered all daytime NEP in the current study. Neural networks are agreed upon by multiple comparison studies as a superior method (Falge et al. 2001, Papale et al. 2006, Moffat et al. 2007) even though usually, differences among methods are often minor (Moffat et al. 2007). Due to availability, we did not use neural networks. These short time scales and the homogenous ecosystem in Ooba et al. (2006) may not reflect long-term trends (Baldocchi et al. 2015) or natural ecosystems. Converting our results into daily estimates does not agree with Ooba et al. (2006); there is much more variance among gap-fill methods. Ooba et al. (2006) only considered four months of data but perhaps over a longer period of time, daily averages can deviate more among gap-fill methods than when compared in only one season. In comparing mean diurnal variation to look-up table and non-linear regressions, yearly estimates of NEP ranged from -200 to +45 g C m^{-2} (mean diurnal variation vs non-linear regression) and -150 to +30 g C m^{-2} (look-up table vs non-linear regression) (Falge et al. 2001). All methods were able to approximate data well with small errors in artificially induced gaps (Falge et al. 2001). Gap-filling errors did not differ among methods and were instead directly proportional to the size of the gap being filled (Falge et al. 2001, Moffat et al. 2007). Since we filled the same dataset with all gap-fill methods, error based on gap size should be the same among all gap-fill methods.

Moffat et al. (2007) concluded that across six forested sites, gap-filling effects were modest and annual NEP among gap-fill methods ranged between ± 25 g C m^{-2} y^{-1}. It was suggested that not much more improvement could occur, and that most of the uncertainty could be attributed to measurement uncertainty as opposed to model uncertainty (Moffat et al. 2007). This range is similar to results in the present study. For example, the largest difference between a gap-fill method and the mean of all four gap-fill methods is between Harvard (4 g C m^{-2} y^{-1}) and the mean 29 g C m^{-2} y^{-1}, totalling 25 g C m^{-2} y^{-1}. This is in comparison to FCRN which is within 7 g C m^{-2} y^{-1} and MPI and Mean Data which were 15 and 19 g C m^{-2} y^{-1} greater than the mean respectively. It is notable as well that all methods that tested significantly different from one another annually differed by at least ±25 g C m^{-2} y^{-1}. In the current study, our error estimate is approximately ±10 g C m^{-2} y^{-1} for these four gap-fill methods at NOBS for the mean annual NEP of 29 g C m^{-2} y^{-1}. A second study found gap-fill method uncertainty to be approximately ± 7 g C m^{-2} y^{-1} but incidences of large gaps were very low (Elbers et al. 2011). Almost every year, gap-fill methods differed by more than 7 g C m^{-2} y^{-1} although not between MPI and Mean Data gap-fill methods. Since all gap-fill methods were used on the exact same dataset, we can attribute the differences to the gap-fill method itself.

2.5.5 Implications for Old-Growth Forests

The uncertainty in the persistence of old-growth forests as sinks is important in the face of a changing climate. In a study encompassing hundreds of old-growth forest datasets encompassing biometry-based net primary production (NPP) and NEP, eddy-covariance measurements and/or chamber-based measurements; it was found that the carbon balance of forests aged 15 to 800 years were usually positive, indicating old-growth forests are on average acting as a net carbon sink (Luyssaert et al. 2008). That study encompassed both boreal and temperate forests; in general boreal forests had lower

NPP than temperate (Luyssaert et al. 2008). Globally, all forests (old and young) are carbon sinks (Pan et al. 2011). In general this is true for NOBS which has remained undisturbed for approximately 160 years. Three of the four gap-fill methods indicated that over 15 years, NOBS was a carbon sink while one estimated neutrality. On average, Luyssaert et al. (2008) estimated that forests over 200 years sequester 2.4 ± 0.8 Mg C ha^{-1} y^{-1}. It has also been found that old-growth forests can continue to accumulate carbon in their soils in southern China at an average rate of 0.035% per year totalling a net increase of 0.95 % over 24 years (Zhou et al. 2006). If NOBS continues to age with no large-scale disturbances (e.g. fire), based on global results, carbon may continue to accumulate as long as the forest is driven by small-scale disturbances (e.g. windthrow) (Amiro et al. 2010) and there is no change in environmental conditions. It is however possible that the rate may slow as the forest ages (Luyssaert et al. 2008).

It is unknown what the response of old-growth forests will be to climate change and a potentially warming atmosphere. In temperate forests (Hember et al. 2012) and boreal forests (Briffa et al. 2008) productivity increases in response to warmer T and increasing atmospheric CO_2 concentrations have been observed. In boreal old-aspen sites, eddy covariance data suggest that carbon sequestration may increase with warming temperatures due to observed conservativeness of R and the fact that photosynthesis will increase with increasing length of growing season (Griffis et al. 2003, Bergeron et al. 2007). Two decades of satellite data from a study by Goetz et al. (2005) suggest that a trend of earlier photosynthesis onset will occur in North American ecosystems. Deciduous forests have a shorter growing season but greater NEP due to higher photosynthetic capacity (Griffis et al. 2003) so would likely benefit even more from longer growing seasons. However, there is the suggestion that increasing temperatures may cause changes in water availability, decomposition dynamics, and soil quality/characteristics that can decrease the carbon accumulating capacity of some northern boreal forests (LaFleur et al. 2010, Beck et al. 2011). This would be of concern if Dunn et al. (2007)'s suggestion that drier conditions increase R at NOBS is true. If water balance changes greatly, especially in the face of increased possibility of droughts (Griffis et al. 2003) there is the possibility of increased carbon loss. Soil moisture has also been found as an important factor, along with temperature, in carbon balance of SOBS forest in Prince Albert National Park, Saskatchewan (Griffis et al. 2003).

The carbon balance of this site is similar to other boreal forests in North America (Griffis et al. 2003, Barr et al. 2007, Bergeron et al. 2007, Krishnan et al. 2008, Ueyama et al. 2014). The Alaskan Old Black Spruce forest site over nine years began as a carbon sink and became a carbon source in the later study years; overall it was near carbon neutral (Ueyama et al. 2014). This shift was associated with warming autumn temperatures. In comparison to other boreal sites (Barr et al. 2007, Krishnan et al. 2008, Trucco et al. 2012) where spring temperature is more important in the annual carbon balance, the Alaskan boreal spruce forests carbon balance was determined by autumnal temperatures (Ueyama et al. 2014). It was suggested that the forest was in the process of undergoing changes related to climate warming (Ueyama et al. 2014) and may be the reason why this site followed a trend differing from other ecosystems in the FLUXNET database (Baldocchi et al. 2008). Increasing temperatures and other changing climatic variables benefited carbon accumulation in boreal old-growth aspen sites in Saskatchewan (Griffis et al. 2003). Dunn et al. (2007) suggested that a warmer, wetter regime may be the reason there was a shift in NOBS carbon balance from a source to a sink during the study period. In northern and western Europe, it is suggested that a few positives will come for boreal forests from increasing CO_2 concentrations, such as greater wood production of trees and overall forest growth on short to moderate time-scales (Lindner et al. 2010). However, these will be counteracted by possible negative effects bringing increasing risk of drought and other disturbances (Lindner et al. 2010). This highlights the importance again of long-term eddy covariance studies in order to determine long-term trends in old-growth forests. In order to create reliable long-term datasets to better compare chronosequences, cross country studies, and inter-continental studies, standardization is pertinent. Processing eddy-covariance data in a similar fashion ensures differences seen between datasets are a product of environmental or climatic differences rather than errors that have biased the data and accumulated errors due to gap-fill methods and allows better inference of changes based on other sites.

2.6 Conclusions

Filling the NOBS 15-year eddy-covariance dataset with four different gap-fill methods led to significant differences in carbon balance both annually and on a monthly basis. Most of these differences were in the same range observed in other studies (Moffat et al. 2007). The goal of the study was not to designate one method to be better than another, but to analyze how different estimates could be, based solely on choice of gap-fill method. In looking at the flux-partitioning routines, estimates were significantly lower when estimated by MPI than by FCRN or Harvard methods. Differences in GEP and R were most apparent in mid-growing season while differences in NEP were most apparent in colder months when flux was small relatively. Over the entire dataset, gap-fill methods did not agree on the ecosystem carbon exchange over time, with FCRN and Harvard methods predicting carbon losses at the beginning of the study but eventually accumulated carbon. When looking at annual NEP, MPI does also reflect this pattern to a weaker extent, but still never demonstrates the forest as a carbon source. Over the entire dataset, MPI and Mean Data methods estimated continual carbon accumulation through the study. This study indicates error due to gap-fill method is large enough to create significant differences among carbon balance estimates although in some cases, this is only apparent over a long-term dataset. Standardization may provide consistent errors for comparisons, whereas technique modification and different techniques allows estimates of uncertainty.

2.7 References

Amiro, B. D., Barr, A. G., Barr, J. G., Black, T. A., Bracho, R., Brown, M., Chen, J., Clark, K. L., Davis, K. J., Desai, A. R., Dore, S., Engel, V., Fuentes, J. D., Goldstein, A. H., Goulden, M. L., Kolb, T. E., Lavigne, M. B., Law, B. E., Margolis, H. A., Martin, T., McCaughey, J. H., Misson, L., Montes-Helu, M., Noormets, A., Randerson, J. T., Starr, G. and Xiao, J. 2010. Ecosystem carbon dioxide fluxes after disturbance in forests of North America. J. Geophys. Res-Biogeo. **115:** G00K02.

Baldocchi, D. 2008. Breathing of the terrestrial biosphere: lessons learned from a global network of carbon dioxide flux measurement systems. Aust. J. Bot. **56:** 1-26.

Baldocchi, D., Sturtevant, C. and Fluxnet Contributors. 2015. Does day and night sampling reduce spurious correlation between canopy photosynthesis and ecosystem respiration? Agric. For. Meteorol. **207:** 117-126.

Barr, A. G., Black, T. A., Hogg, E. H., Kljun, N., Morgenstern, K. and Nesic, Z. 2004. Inter-annual variability in the leaf area index of a boreal aspen-hazelnut forest in relation to net ecosystem production. Agric. For. Meteorol. **126:** 237-255.

Barr, A. G., Griffis, T. J., Black, T. A., Lee, X., Staebler, R. M., Fuentes, J. D., Chen, Z. and Morgenstern, K. 2002. Comparing the carbon budgets of boreal and temperate deciduous forest stands. Can. J. Forest Res. **32:** 813-822.

Barr, A. G., Black, T. A., Hogg, E. H., Griffis, T. J., Morgenstern, K., Kljun, N., Theede, A. and Nesic, Z. 2007. Climatic controls on the carbon and water balances of a boreal aspen forest, 1994-2003. Global Change Biol. **13:** 561-576.

Beck, P. S. A., Juday, G. P., Alix, C., Barber, V. A., Winslow, S. E., Sousa, E. E., Heiser, P., Herriges, J. D. and Goetz, S. J. 2011. Changes in forest productivity across Alaska consistent with biome shift. Ecol. Lett. **14:** 373-379.

Bergeron, O., Margolis, H. A., Black, T. A., Coursolle, C., Dunn, A. L., Barr, A. G. and Wofsy, S. C. 2007. Comparison of carbon dioxide fluxes over three boreal black spruce forests in Canada. Global Change Biol. **13:** 89-107.

Black, T. A., Gaumont-Guay, D., Jassal, R. S., Amiro, B. D., Jarvis, P. G., Gower, S. T., Kelliher, F. M., Dunn, A. and Wofsy, S. C. 2005. Measurement of CO_2 exchange between Boreal forest and the atmosphere. SEB Exp. Biol. Ser. 2005: 151-85.

Bonan, G. B., Lawrence, P. J., Oleson, K. W., Levis, S., Jung, M., Reichstein, M., Lawrence, D. M. and Swenson, S. C. 2011. Improving canopy processes in the Community Land Model version 4 (CLM4) using global flux fields empirically inferred from FLUXNET data. J. Geophys. Res-Biogeo. **116:** G02014.

Brandt, J. P. 2009. The extent of the North American boreal zone. Env. Rev. **17:** 101-161.

Briffa, K. R., Shishov, V. V., Melvin, T. M., Vaganov, E. A., Grudd, H., Hantemirov, R. M., Eronen, M. and Naurzbaev, M. M. 2008. Trends in recent temperature and radial tree growth spanning 2000 years across northwest Eurasia. Philos. T. Roy. Soc. B. **363:** 2271-2284.

Coursolle, C., Margolis, H. A., Giasson, M. A., Bernier, P. Y., Amiro, B. D., Arain, M. A., Barr, A. G., Black, T. A., Goulden, M. L., McCaughey, J. H., Chen, J. M., Dunn, A. L., Grant, R. F. and Lafleur, P. M. 2012. Influence of stand age on the magnitude and seasonality of carbon fluxes in Canadian forests. Agric. For. Meteorol. **165:** 136-148.

Desai, A. R., Richardson, A. D., Moffat, A. M., Kattge, J., Hollinger, D. Y., Barr, A., Falge, E., Noormets, A., Papale, D., Reichstein, M. and Stauch, V. J. 2008. Cross-site evaluation of eddy covariance GPP and RE decomposition techniques. Agric. For. Meteorol. **148:** 821-838.

Dragomir, C. M., Klaassen, W., Voiculescu, M., Georgescu, L. P. and van der Laan, S. 2012. Estimating annual CO_2 flux for Lutjewad Station using three different gap-filling techniques. Scientific World Journal **2012:** 1-10.

Dragoni, D., Schmid, H. P., Grimmond, C. S. B. and Loescher, H. W. 2007. Uncertainty of annual net ecosystem productivity estimated using eddy covariance flux measurements. J. Geophys. Res.-Atmos. **112:** D17102.

Dunn, A. L., Barford, C. C., Wofsy, S. C., Goulden, M. L. and Daube, B. C. 2007. A long-term record of carbon exchange in a boreal black spruce forest: means, responses to interannual variability, and decadal trends. Global Change Biol. **13:** 577-590.

Elbers, J. A., Jacobs, C. M. J., Kruijt, B., Jans, W. W. P. and Moors, E. J. 2011. Assessing the uncertainty of estimated annual totals of net ecosystem productivity: A practical approach applied to a mid latitude temperate pine forest. Agric. For. Meteorol. **151:** 1823-1830.

Falge, E., Baldocchi, D., Olson, R., Anthoni, P., Aubinet, M., Bernhofer, C., Burba, G., Ceulemans, R., Clement, R., Dolman, H., Granier, A., Gross, P., Grunwald, T., Hollinger, D., Jensen, N. O., Katul, G., Keronen, P., Kowalski, A., Lai, C. T., Law, B. E., Meyers, T., Moncrieff, H., Moors, E., Munger, J. W., Pilegaard, K., Rannik, U., Rebmann, C., Suyker, A., Tenhunen, J., Tu, K., Verma, S., Vesala, T., Wilson, K. and Wofsy, S. 2001. Gap filling strategies for defensible annual sums of net ecosystem exchange. Agric. For. Meteorol. **107:** 43-69.

Falge, E., Baldocchi, D., Tenhunen, J., Aubinet, M., Bakwin, P., Berbigier, P., Bernhofer, C., Burba, G., Clement, R., Davis, K. J., Elbers, J. A., Goldstein, A. H.,

Grelle, A., Granier, A., Guomundsson, J., Hollinger, D., Kowalski, A. S., Katul, G., Law, B. E., Malhi, Y., Meyers, T., Monson, R. K., Munger, J. W., Oechel, W., Paw, K. T., Pilegaard, K., Rannik, U., Rebmann, C., Suyker, A., Valentini, R., Wilson, K. and Wofsy, S. 2002. Seasonality of ecosystem respiration and gross primary production as derived from FLUXNET measurements. Agric. For. Meteorol. **113:** 53-74.

Fu, D., Chen, B., Zhang, H., Wang, J., Black, T. A., Amiro, B. D., Bohrer, G., Bolstad, P., Coulter, R., Rahman, A. F., Dunn, A., McCaughey, J. H., Meyers, T. and Verma, S. 2014. Estimating landscape net ecosystem exchange at high spatial-temporal resolution based on Landsat data, an improved upscaling model framework, and eddy covariance flux measurements. Remote Sens. Environ. **141:** 90-104.

Goetz, S. J., Bunn, A. G., Fiske, G. J. and Houghton, R. A. 2005. Satellite-observed photosynthetic trends across boreal North America associated with climate and fire disturbance. Proc. Natl. Acad. Sci. U. S. A. **102:** 13521-13525.

Gorham, E. 1991. Northern Peatlands - Role in the Carbon-Cycle and Probable Responses to Climatic Warming. Ecol. Appl. **1:** 182-195.

Goulden, M. L., McMillan, A. M. S., Winston, G. C., Rocha, A. V., Manies, K. L., Harden, J. W. and Bond-Lamberty, B. P. 2011. Patterns of NPP, GPP, respiration, and NEP during boreal forest succession. Global Change Biol. **17:** 855-871.

Goulden, M. L., Wofsy, S. C., Harden, J. W., Trumbore, S. E., Crill, P. M., Gower, S. T., Fries, T., Daube, B. C., Fan, S. M., Sutton, D. J., Bazzaz, A. and Munger, J. W. 1998. Sensitivity of boreal forest carbon balance to soil thaw. Science **279:** 214-217.

Goulden, M. L., Winston, G. C., McMillan, A. M. S., Litvak, M. E., Read, E. L., Rocha, A. V. and Elliot, J. R. 2006. An eddy covariance mesonet to measure the effect of forest age on land-atmosphere exchange. Global Change Biol. **12:** 2146-2162.

Gower, S. T., Vogel, J. G., Norman, J. M., Kucharik, C. J., Steele, S. J. and Stow, T. K. 1997. Carbon distribution and aboveground net primary production in aspen, jack pine, and black spruce stands in Saskatchewan and Manitoba, Canada. J. Geophys. Res.-Atmos. **102:** 29029-29041.

Griffis, T. J., Black, T. A., Morgenstern, K., Barr, A. G., Nesic, Z., Drewitt, G. B., Gaumont-Guay, D. and McCaughey, J. H. 2003. Ecophysiological controls on the carbon balances of three southern boreal forests. Agric. For. Meteorol. **117:** 53-71.

Harden, J. W., O'Neill, K. P., Trumbore, S. E., Veldhuis, H. and Stocks, B. J. 1997. Moss and soil contributions to the annual net carbon flux of a maturing boreal forest. J. Geophys. Res.-Atmos. **102:** 28805-28816.

Heinsch, F. A., Zhao, M., Running, S. W., Kimball, J. S., Nemani, R. R., Davis, K. J., Bolstad, P. V., Cook, B. D., Desai, A. R., Ricciuto, D. M., Law, B. E., Oechel, W. C.,

Kwon, H., Luo, H., Wofsy, S. C., Dunn, A. L., Munger, J. W., Baldocchi, D. D., Xu, L., Hollinger, D. Y., Richardson, A. D., Stoy, P. C., Siqueira, M. B. S., Monson, R. K., Burns, S. P. and Flanagan, L. B. 2006. Evaluation of remote sensing based terrestrial productivity from MODIS using regional tower eddy flux network observations. IEEE Trans. Geosci. Remote Sens. **44:** 1908-1925.

Hember, R. A., Kurz, W. A., Metsaranta, J. M., Black, T. A., Guy, R. D. and Coops, N. C. 2012. Accelerating regrowth of temperate-maritime forests due to environmental change. Global Change Biol. **18:** 2026-2040.

Hilton, T. W., Davis, K. J. and Keller, K. 2014. Evaluating terrestrial CO_2 flux diagnoses and uncertainties from a simple land surface model and its residuals. Biogeosciences **11:** 217-235.

Horn, J. E. and Schulz, K. 2011. Identification of a general light use efficiency model for gross primary production. Biogeosciences **8:** 999-1021.

Hui, D. F., Wan, S. Q., Su, B., Katul, G., Monson, R. and Luo, Y. Q. 2004. Gap-filling missing data in eddy covariance measurements using multiple imputation (MI) for annual estimations. Agric. For. Meteorol. **121:** 93-111.

Krishnan, P., Black, T. A., Barr, A. G., Grant, N. J., Gaumont-Guay, D. and Nesic, Z. 2008. Factors controlling the interannual variability in the carbon balance of a southern boreal black spruce forest. J. Geophys. Res.-Atmos. **113:** D09109.

Kurz, W. A., Shaw, C. H., Boisvenue, C., Stinson, G., Metsaranta, J., Leckie, D., Dyk, A., Smyth, C. and Neilson, E. T. 2013. Carbon in Canada's boreal forest - A synthesis. Env. Rev. **21:** 260-292.

Lafleur, B., Pare, D., Munson, A. D. and Bergeron, Y. 2010. Response of northeastern North American forests to climate change: Will soil conditions constrain tree species migration? Env. Rev. **18:** 279-289.

Lasslop, G., Migliavacca, M., Bohrer, G., Reichstein, M., Bahn, M., Ibrom, A., Jacobs, C., Kolari, P., Papale, D., Vesala, T., Wohlfahrt, G. and Cescatti, A. 2012. On the choice of the driving temperature for eddy-covariance carbon dioxide flux partitioning. Biogeosciences **9:** 5243-5259.

Lindner, M., Maroschek, M., Netherer, S., Kremer, A., Barbati, A., Garcia-Gonzalo, J., Seidl, R., Delzon, S., Corona, P., Kolstrom, M., Lexer, M. J. and Marchetti, M. 2010. Climate change impacts, adaptive capacity, and vulnerability of European forest ecosystems. For. Ecol. Manage. **259:** 698-709.

Lloyd, J. and Taylor, J. A. 1994. On the temperature-dependence of soil respiration. Funct. Ecol. **8:** 315-323.

Luyssaert, S., Schulze, E. D., Boerner, A., Knohl, A., Hessenmoeller, D., Law, B. E., Ciais, P. and Grace, J. 2008. Old-growth forests as global carbon sinks. Nature **455:** 213-215.

Mahecha, M. D., Reichstein, M., Carvalhais, N., Lasslop, G., Lange, H., Seneviratne, S. I., Vargas, R., Ammann, C., Arain, M. A., Cescatti, A., Janssens, I. A., Migliavacca, M., Montagnani, L. and Richardson, A. D. 2010. Global convergence in the temperature sensitivity of respiration at ecosystem level. Science **329:** 838-840.

Moffat, A. M., Papale, D., Reichstein, M., Hollinger, D. Y., Richardson, A. D., Barr, A. G., Beckstein, C., Braswell, B. H., Churkina, G., Desai, A. R., Falge, E., Gove, J. H., Heimann, M., Hui, D., Jarvis, A. J., Kattge, J., Noormets, A. and Stauch, V. J. 2007. Comprehensive comparison of gap-filling techniques for eddy covariance net carbon fluxes. Agric. For. Meteorol. **147:** 209-232.

Morgenstern, K., Black, T. A., Humphreys, E. R., Griffis, T. J., Drewitt, G. B., Cai, T. B., Nesic, Z., Spittlehouse, D. L. and Livingston, N. J. 2004. Sensitivity and uncertainty of the carbon balance of a Pacific Northwest Douglas-fir forest during an El Nino La Nina cycle. Agric. For. Meteorol. **123:** 201-219.

Ooba, M., Hirano, T., Mogami, J., Hirata, R. and Fujinuma, Y. 2006. Comparisons of gap-filling methods for carbon flux dataset: A combination of a genetic algorithm and an artificial neural network. Ecol. Model. **198:** 473-486.

Pan, Y., Birdsey, R. A., Fang, J., Houghton, R., Kauppi, P. E., Kurz, W. A., Phillips, O. L., Shvidenko, A., Lewis, S. L., Canadell, J. G., Ciais, P., Jackson, R. B., Pacala, S. W., McGuire, A. D., Piao, S., Rautiainen, A., Sitch, S. and Hayes, D. 2011. A large and persistent carbon sink in the world's forests. Science **333:** 988-993.

Papale, D., Reichstein, M., Aubinet, M., Canfora, E., Bernhofer, C., Kutsch, W., Longdoz, B., Rambal, S., Valentini, R., Vesala, T. and Yakir, D. 2006. Towards a standardized processing of Net Ecosystem Exchange measured with eddy covariance technique: algorithms and uncertainty estimation. Biogeosciences **3:** 571-583.

Reichstein, M., Falge, E., Baldocchi, D., Papale, D., Aubinet, M., Berbigier, P., Bernhofer, C., Buchmann, N., Gilmanov, T., Granier, A., Grunwald, T., Havrankova, K., Ilvesniemi, H., Janous, D., Knohl, A., Laurila, T., Lohila, A., Loustau, D., Matteucci, G., Meyers, T., Miglietta, F., Ourcival, J. M., Pumpanen, J., Rambal, S., Rotenberg, E., Sanz, M., Tenhunen, J., Seufert, G., Vaccari, F., Vesala, T., Yakir, D. and Valentini, R. 2005. On the separation of net ecosystem exchange into assimilation and ecosystem respiration: review and improved algorithm. Global Change Biol. **11:** 1424-1439.

Reichstein, M., Tenhunen, J. D., Roupsard, O., Ourcival, J. M., Rambal, S., Dore, S. and Valentini, R. 2002. Ecosystem respiration in two Mediterranean evergreen Holm Oak forests: drought effects and decomposition dynamics. Funct. Ecol. **16:** 27-39.

Rocha, A. V., Goulden, M. L., Dunn, A. L. and Wofsy, S. C. 2006. On linking interannual tree ring variability with observations of whole-forest CO_2 flux. Global Change Biol. **12:** 1378-1389.

Schwalm, C. R., Williams, C. A., Schaefer, K., Arneth, A., Bonal, D., Buchmann, N., Chen, J., Law, B. E., Lindroth, A., Luyssaert, S., Reichstein, M. and Richardson, A. D. 2010. Assimilation exceeds respiration sensitivity to drought: A FLUXNET synthesis. Global Change Biol. **16:** 657-670.

Sellers, P., Hall, F., Margolis, H., Kelly, B., Baldocchi, D., den Hartog, G., Cihlar, J., Ryan, M. G., Goodison, B., Crill, P., Ranson, K. J., Lettenmaier, D. and Wickland, D. E. 1995. The Boreal Ecosystem-Atmosphere Study (BOREAS) - an overview and early results from the 1994 field year. Bull. Am. Meteorol. Soc. **76:** 1549-1577.

Soja, A. J., Tchebakova, N. M., French, N. H. F., Flannigan, M. D., Shugart, H. H., Stocks, B. J., Sukhinin, A. I., Parfenova, E. I., Chapin, F. S.,III and Stackhouse, P. W.,Jr. 2007. Climate-induced boreal forest change: Predictions versus current observations. Global Planet. Change **56:** 274-296.

Stoy, P. C., Katul, G. G., Siqueira, M. B. S., Juang, J., Novick, K. A., Uebelherr, J. M. and Oren, R. 2006. An evaluation of models for partitioning eddy covariance-measured net ecosystem exchange into photosynthesis and respiration. Agric. For. Meteorol. **141:** 2-18.

Stoy, P. C., Trowbridge, A. M. and Bauerle, W. L. 2014. Controls on seasonal patterns of maximum ecosystem carbon uptake and canopy-scale photosynthetic light response: contributions from both temperature and photoperiod. Photosynthesis Res. **119:** 49-64.

Subke, J. and Bahn, M. 2010. On the 'temperature sensitivity' of soil respiration: Can we use the immeasurable to predict the unknown? Soil Biol. Biochem. **42:** 1653-1656.

Suni, T., Rinne, J., Reissell, A., Altimir, N., Keronen, P., Rannik, U., Dal Maso, M., Kulmala, M. and Vesala, T. 2003. Long-term measurements of surface fluxes above a Scots pine forest in Hyytiala, southern Finland, 1996-2001. Boreal Environ. Res. **8:** 287-301.

Trucco, C., Schuur, E. A. G., Natali, S. M., Belshe, E. F., Bracho, R. and Vogel, J. 2012. Seven-year trends of CO_2 exchange in a tundra ecosystem affected by long-term permafrost thaw. J. Geophys. Res-Biogeo. **117:** G02031.

Trumbore, S. E. and Harden, J. W. 1997. Accumulation and turnover of carbon in organic and mineral soils of the BOREAS northern study area. J. Geophys. Res.-Atmos. **102:** 28817-28830.

Turner, D. P., Ritts, W. D., Zhao, M., Kurc, S. A., Dunn, A. L., Wofsy, S. C., Small, E. E. and Running, S. W. 2006. Assessing interannual variation in MODIS-based estimates of gross primary production. IEEE Trans. Geosci. Remote Sens. **44:** 1899-1907.

Ueyama, M., Iwata, H. and Harazono, Y. 2014. Autumn warming reduces the CO2 sink of a black spruce forest in interior Alaska based on a nine-year eddy covariance measurement. Global Change Biol. **20:** 1161-1173.

Valentini, R., Dore, S., Marchi, G., Mollicone, D., Panfyorov, M., Rebmann, C., Kolle, O. and Schulze, E. D. 2000. Carbon and water exchanges of two contrasting central Siberia landscape types: regenerating forest and bog. Funct. Ecol. **14:** 87-96.

Zhou, G., Liu, S., Li, Z., Zhang, D., Tang, X., Zhou, C., Yan, J. and Mo, J. 2006. Old-growth forests can accumulate carbon in soils. Science **314:** 1417-1417.

3. OVERALL SYNTHESIS

3.1 Summary

I filled a 15-year eddy covariance dataset measured at NOBS forest near Thompson, in central Manitoba, Canada using four different gap-fill methods: FCRN, Harvard, MPI, and Mean Data method. The flux tower at NOBS was established as part of NASAs BOREAS atmosphere-biosphere monitoring project. Measurements were taken spanning from 1994 to 2008. The eddy covariance dataset was available from the FLUXNET database online (http://fluxnet.ornl.gov/). We then filled the dataset with the four different gap-fill methods and compared annual cumulative NEP, monthly cumulative NEP, and annual flux-partitioning components (GEP and R) among gap-fill methods in an attempt to determine if there were any differences. Our objectives were to determine whether choice of gap-fill method significantly affected the results of annual NEP, and to see if we could determine why results differed based on GEP and R.

Annually, NEP was estimated as low as 4.5 g C m^{-2} y^{-1} by Harvard gap-fill, 21.9 g C m^{-2} y^{-1} by FCRN gap-fill, and similarly between MPI and Mean Data gap-fill methods as 43.5 and 48.0 g C m^{-2} y^{-1} respectively (averaged from 1995-2006). The average among all gap-fill methods was 29 ± 10 g C m^{-2} y^{-1}. We were able to determine that annually, both MPI and Mean Data gap-fill methods did not significantly differ from one another, but both were significantly more positive than a hypothetical carbon neutral forest. Harvard gap-fill and FCRN did not differ significantly from one another but were both more positive that the hypothetical carbon neutral forest. Mean Data and MPI differed significantly from Harvard gap-fill; with Mean Data and MPI being significantly more positive.

Cumulative carbon balances for NOBS differed depending on gap-fill method. Both MPI and Mean Data gap-fill methods predicted carbon accumulation across the whole dataset reaching approximately 695 g C m^{-2} and 768 g C m^{-2} respectively over 15 years. Harvard gap-fill estimated a period of time when NOBS acted as a carbon source early in the study. FCRN predicted the same pattern of carbon balance over the 15-year study period but to a less exaggerated extent. When looking at NEP annual data, MPI also demonstrated this pattern to an even weaker extent. Over the whole data set, FCRN and Harvard-filled data reached a cumulative NEP of 373 g C m^{-2} (2008) and -33 g C m^{-2} (2006) respectively.

Although comparisons encompassing a broad range of forested sites and gap-fill methods have indicated that the choice of gap-fill method may not be important on an annual basis (Falge et al. 2001, Moffat et al. 2007), it seems that over time, these smaller errors can compound and create large differences in cumulative NEP over longer periods. Flux-partitioning does not appear to explain all the differences seen among gap-fill methods. Data partitioned by MPI (R and GEP) was always significantly lower than Harvard and FCRN. In one year of lower than average T and precipitation (2004), some divergences were seen between Harvard and FCRN GEP and R. Some disagreement between these two methods was seen in winter estimates of R. This could explain some of the differences seen in cumulative NEP.

Here we conclude choice of gap-fill method may only create modest differences in annual NEP but over time, errors can compound creating great differences in cumulative NEP over a long-term eddy-covariance dataset. Differences in cumulative NEP may simply be caused by method of calculation as no specific methodological cause could be determined. The current study did not attempt to appoint one gap-fill method as superior to another, but rather attempted to characterize the range of values that can be obtained in using such gap-fill methods. This study provides further indications that processing eddy-covariance data should be standardized in order to facilitate synthesis activities and inter-comparisons among sites and ecosystem types. Additionally, reporting errors and taking into account gap-fill method errors could help alleviate some of these issues and are important in comparison and synthesis studies using eddy-covariance measurements.

3.2 Significance

Being able to accurately characterize the carbon balance of any forest is of great importance in global carbon balances and budgets, as well as forest management and carbon emission mitigation strategies. Understanding the carbon dynamics in the event of droughts, temperature increases and changes in other environmental drivers such as growing season length provides a lot of predictive power and insight into important driving variables. Characterization of forests in the boreal biome is of particular importance due to the vulnerability of these areas in response to a changing climate. This area is one of predicted rapid change (Soja et al. 2007). The carbon balance and stability of old-growth forests is uncertain over a long period of time in the future, especially in the context of a changing climate. It is important not only to monitor these ecosystems, but also to ensure the data we obtain are an accurate representation of what is actually going on with carbon flux. This includes better knowledge on the uncertainty involved in gap-filling and the processing used in gap-filling.

Since one of the main limitations of eddy-covariance studies is the need for sufficient turbulent motions, many gaps in data measurements occur. This occurs mainly during nighttime measurements. There are gaps that occur through the dataset as well because of system malfunction, component failure, or the need to exclude measurements from certain areas. Gap-filling is extremely useful in order to obtain a complete dataset and is a step towards having the ability to model 100% of data based solely on associated meteorological data. This would allow the omission of expensive measurement set-ups and the maintenance of long-term tower sites. In the meantime, it is important these methods become as accurate as possible, and that conclusions drawn from eddy-covariance are based on observed differences in ecosystem processes and not based on errors due to different processing methods. This study further emphasizes the importance of standardizing gap-fill methods by demonstrating the wide variety of estimates possible within one single site. Comparing sites using the same gap-fill method provides the capability of capturing the relative differences between sites at the very least. Comparing between sites filled with different gap-fill methods will lead to conclusions clouded with uncertainty due to gap-fill method or error from other sources if those are not controlled for either. Based on the current study the range in annual NEP values within a site can be as large as 45 g C m^{-2} y^{-1} based on gap-fill methods. This is similar to the range of values ± 25 g C m^{-2} y^{-1} observed in studies encompassing numerous forest sites (Moffat et al. 2007). The actual error we determined for these four gap-fill methods at NOBS was ±10 g C m^{-2} y^{-1}. Accurate error quantification of the processing techniques will also facilitate synthesis efforts and provide much needed improvements in comparisons making chronosequences more reliable and more informative for studying effects of climate change or effect of aging on forests.

3.3 Improvements

An improvement that could have been accomplished in our study was having reliable soil temperature data available. Unfortunately, due to the longevity of the study, over time, soil temperature calibration shifted making any measurements after the first few years unusable. It is known in the flux data community that filling data based on air T or soil T will yield different flux estimates (Lasslop et al. 2012) and it is not known which is better (Subke and Bahn 2010). Air T has much larger diurnal variation than soil T and may trigger seasonal changes sooner than soil T. Many plant processes are driven more by soil T than air T. It has been suggested to test both air and soil T and determine which best explains the variation in site R (Lasslop et al. 2012). We did not perform such a test and used the most complete and accurate data available.

Much of the data in this study was provided to us by external sources. This includes the entire Harvard gap-fill dataset which was limited to only 13 years of data, since gap-filling was only available until the end of 2006. Having full access to the computer code used to fill this dataset would have allowed us to fill the remaining two years of dataset and adjust quality control to be in sync with those used in the three other gap-fill methods. Similarly, we did not have access to the code used for the MPI gap-fill as it was submitted through an online gap-filling tool. This left some uncertainty in algorithms used and any queries had to be rectified through email with the scientist in charge.

It was unfortunate that the final two years of the study were of greater gap percentages due to many generator issues causing problems with both primary and backup power. This downfall led to the highest amount of data gaps in 2007 and 2008. Larger gaps create more issues in gap-filling and long periods of no data measurements were present (Falge et al. 2001), including at times during peak growing season. Data during peak growing season is more difficult to model and NEP, R and GEP are proportionally larger than winter and nighttime data which can lead to large errors in estimates. It is estimated that a week-long gap in spring can add as much as ±30 g C m^{-2} y^{-1} in error (Richardson and Hollinger 2007).

This study would benefit from an even longer dataset. Evidently, this is an improvement that can be rectified if a different flux site is used with longer tower presence since NOBS was decommissioned in 2008. This would allow comparison of gap-fill methods during an even longer dataset, but more importantly, this would encompass a greater variety of environmental conditions, including drought, extreme weather events, and perhaps some sort of natural disturbance like fire or insect infestation. A longer dataset would also allow comparison of fluxes over time if it were observed that temperature increased significantly during the study period or if there were significant changes in precipitation regimes.

Something that could have been considered in the current study was the use of an artificial neural network as a gap-fill method. This has been singled out as perhaps the superior gap-fill method by a number of studies (Falge et al. 2001, Ooba et al. 2006, Papale et al. 2006, Moffat et al. 2007, Desai et al. 2008), and we were not able to compare it to these other methods. The neural network method is a non-linear regression technique involving the presentation of input data (PAR and/or air T) and output data (NEE) to a series of nodes and weights (the network) representing the regression parameters (Papale and Valentini 2003, Moffat et al. 2007, Desai et al. 2008). By re-running these data through the network, it mimics neural learning, and can weight some connections more than others and vice versa, in order to yield the most accurate estimate. This would reflect how different drivers may affect NEE differently than others in certain instances. The drivers and their effects on NEE are then mapped into the network and then it is used to fill the gaps in NEE (Moffat et al. 2007). There are different types of neural networks including some using specific algorithms (Braswell et al. 2005) or using specific training datasets that are then run through multiple neural networks with different architectures, in order to receive a final averaged NEE value (Papale and Valentini 2003). Being able to test a more complex and a forerunner for most accurate gap-fill method could have benefitted this study by having a completely different process for filling gaps. For example, Mean Data and MPI both use a form of averaging data and do not use drivers for filling NEE specifically. The main reason for not using neural networks was coding availability and by using this method, weighting and associated equations would be different if applied to other sites.

The only input variables used in gap-filling as environmental drivers are air T and PAR. This requires the assumption that R and GEP are greatly linked to changes in PAR and air T. Any variability in ecosystem function that is not captured by these inputs will not be reflected in the estimated R or GEP and in turn NEP (Desai et al. 2008). The issue as well is that not all ecosystem types will respond to these driving variables similarly, some will respond more strongly or weakly to these inputs depending on climate zone. Neural networks have attempted to alleviate some of this by allowing the network to be trained and weight regression parameters differently, so may be more applicable across ecosystem types (Desai et al. 2008). In an attempt to increase understanding of climate and biological drivers on carbon exchange at different time scales between different forests and the atmosphere; a process-based model (CASTANEA) was used to simulate carbon flux over forests (Delpierre et al. 2012). The controls on boreal forest gross primary production (GPP) were not weighted strongly towards either climatic or biological (Delpierre et al. 2012). Opposing predictions (Reichstein et al. 2007), temperature was not a prominent driver of GPP in comparison to incoming radiation and maximum leaf area (Delpierre et al. 2012). The driving patterns between NEP and GPP had many similarities so it was determined a major control of net C balance was variations in photosynthesis (Delpierre et al. 2012); This was contradicted slightly by the boreal forest since these endure strong thermal limitations (Suni et al. 2003). R was more dependent on soil water (explaining 40% of the variance) emphasized by the few drought years in the study period (Delpierre et al. 2012). This was also observed in the NOBS study where water table explained much of the variability in R (Dunn et al. 2007). In a boreal-temperate transition forest, determined to be a moderate carbon sink, PAR and soil T were the dominant environmental drivers for the long-term trends in NEP (Froelich et al. 2015). Froelich et al. (2015) suggest large interannual variability in NEP merits the implementation of long-term studies. On a shorter time-scale (weeks-months), air T and precipitation were also significant drivers. In boreal bogs, NEP variability was linked mostly to changes in water table depth (Strachan et al. 2016), similar to NOBS (Dunn et al. 2007) and SOBS (Krishnan et al. 2008). Not all important driving variables can be accounted for in simpler gap-fill methods, but changes can be inferred if reflected in T air and/or PAR.

3.4 Future Studies

The most useful studies that could stem from this one would be performing similar analysis on forests or ecosystems in different biomes in order to determine whether gap-fill method choice is site dependent or whether gap-filling is dependent on the range of temperatures or environmental conditions experienced by the site. It is important that NEP values or flux component values be reported with a value of uncertainty and more studies of this nature may be able to decipher some sort of value for specific gap-fill methods. This could aid in cross-site comparison allowing the researcher to account for errors added to the data based on gap-fill method. It would be beneficial to compare gap-fill method effects on flux datasets measured from disturbed sites or during recovery of a disturbance or estimates of carbon loss due to a disturbance since fluxes can change rapidly during these times. In general what would be more important in the flux community is the continued support of flux tower measurements and increasing the length of flux datasets so that more long-term studies can be used to learn more about carbon balance with our increasing knowledge of errors in flux datasets. Up until now, many efforts to determine gap-fill error have used annual flux only, but it would be important to understand the effects over long datasets as they become more common and towers are in existence for longer periods of time.

3.5 References

Delpierre, N., Soudani, K., Francois, C., Le Maire, G., Bernhofer, C., Kutsch, W., Misson, L., Rambal, S., Vesala, T. and Dufrene, E. 2012. Quantifying the influence of climate and biological drivers on the interannual variability of carbon exchanges in European forests through process-based modelling. Agric. For. Meteorol. **154:** 99-112.

Desai, A. R., Richardson, A. D., Moffat, A. M., Kattge, J., Hollinger, D. Y., Barr, A., Falge, E., Noormets, A., Papale, D., Reichstein, M. and Stauch, V. J. 2008. Cross-site evaluation of eddy covariance GPP and RE decomposition techniques. Agric. For. Meteorol. **148:** 821-838.

Falge, E., Baldocchi, D., Olson, R., Anthoni, P., Aubinet, M., Bernhofer, C., Burba, G., Ceulemans, R., Clement, R., Dolman, H., Granier, A., Gross, P., Grunwald, T., Hollinger, D., Jensen, N. O., Katul, G., Keronen, P., Kowalski, A., Lai, C. T., Law, B. E., Meyers, T., Moncrieff, H., Moors, E., Munger, J. W., Pilegaard, K., Rannik, U., Rebmann, C., Suyker, A., Tenhunen, J., Tu, K., Verma, S., Vesala, T., Wilson, K. and Wofsy, S. 2001. Gap filling strategies for defensible annual sums of net ecosystem exchange. Agric. For. Meteorol. **107:** 43-69.

Froelich, N., Croft, H., Chen, J. M., Gonsamo, A. and Staebler, R. M. 2015. Trends of carbon fluxes and climate over a mixed temperate-boreal transition forest in southern Ontario, Canada. Agric. For. Meteorol. **211:** 72-84.

Krishnan, P., Black, T. A., Barr, A. G., Grant, N. J., Gaumont-Guay, D. and Nesic, Z. 2008. Factors controlling the interannual variability in the carbon balance of a southern boreal black spruce forest. J. Geophys. Res.-Atmos. **113:** D09109.

Lasslop, G., Migliavacca, M., Bohrer, G., Reichstein, M., Bahn, M., Ibrom, A., Jacobs, C., Kolari, P., Papale, D., Vesala, T., Wohlfahrt, G. and Cescatti, A. 2012. On the choice of the driving temperature for eddy-covariance carbon dioxide flux partitioning. Biogeosciences **9:** 5243-5259.

Moffat, A. M., Papale, D., Reichstein, M., Hollinger, D. Y., Richardson, A. D., Barr, A. G., Beckstein, C., Braswell, B. H., Churkina, G., Desai, A. R., Falge, E., Gove, J. H., Heimann, M., Hui, D., Jarvis, A. J., Kattge, J., Noormets, A. and Stauch, V. J. 2007. Comprehensive comparison of gap-filling techniques for eddy covariance net carbon fluxes. Agric. For. Meteorol. **147:** 209-232.

Ooba, M., Hirano, T., Mogami, J., Hirata, R. and Fujinuma, Y. 2006. Comparisons of gap-filling methods for carbon flux dataset: A combination of a genetic algorithm and an artificial neural network. Ecol. Model. **198:** 473-486.

Papale, D., Reichstein, M., Aubinet, M., Canfora, E., Bernhofer, C., Kutsch, W., Longdoz, B., Rambal, S., Valentini, R., Vesala, T. and Yakir, D. 2006. Towards a standardized processing of Net Ecosystem Exchange measured with eddy covariance technique: algorithms and uncertainty estimation. Biogeosciences **3:** 571-583.

Papale, D. and Valentini, R. 2003. A new assessment of European forests carbon exchanges by eddy fluxes and artificial neural network spatialization. Global Change Biol. **9:** 525-535.

Richardson, A.D., and D.Y. Hollinger. 2007. A method to estimate the additional uncertainty in gap-filled NEE resulting from long gaps in the CO_2 flux record. Agric. For. Meteorol. **147:** 199-208.

Reichstein, M., Papale, D., Valentini, R., Aubinet, M., Bernhofer, C., Knohl, A., Laurila, T., Lindroth, A., Moors, E., Pilegaard, K. and Seufert, G. 2007. Determinants of terrestrial ecosystem carbon balance inferred from European eddy covariance flux sites. Geophys. Res. Lett. **34:** L01402.

Strachan, I. B., Pelletier, L. and Bonneville, M. 2016. Inter-annual variability in water table depth controls net ecosystem carbon dioxide exchange in a boreal bog. Biogeochemistry **127:** 99-111.

Subke, J. and Bahn, M. 2010. On the 'temperature sensitivity' of soil respiration: Can we use the immeasurable to predict the unknown? Soil Biol. Biochem. **42:** 1653-1656.

Suni, T., Berninger, F., Markkanen, T., Keronen, P., Rannik, U. and Vesala, T. 2003. Interannual variability and timing of growing-season CO_2 exchange in a boreal forest. J. Geophys. Res.-Atmos. **108:** 4265.