

## 8. Discussion

The primary aim of this thesis was to examine and reduce the sources of uncertainty for the estimation of regional to catchment scale carbon budgets. We initially partitioned this task into the influence of the parameterisation of the exchange surface (primarily quantified by LAI), and uncertainties in meteorological driving variables. An investigation of these issues was followed by an analysis of the relative magnitude of uncertainties in the C budget attributable to parameterisation and driver errors respectively. We demonstrated that the dominant source of uncertainty in the final analysis of the C budget was land surface parameterisation, although issues of bias in driver upscaling remain to be resolved.

Land surface parameterisation must be improved to make reliable estimates of C budgets on a regional scale. In Chapter 3 we report persistence of the functional form and approximate error magnitude for LAI NDVI relationships at multiple scales (Williams et al., 2008). However, despite this promising result we demonstrate that it is insufficient to rely on EO derived vegetation indices to provide land surface parameterisations, with weak but significant relationships between key ecosystem variables (LAI) and NDVI (chapters 3 and 4).

On a global scale, DGVMs tend to parameterise the land surface based on a vegetation classification of plant functional types (PFTs) (e.g. Woodward et al., 1995). Results from chapter 4 seem to bear out the validity of this approach. However, the clear variation within PFTs evident in chapter 4 indicates that community dependent topographic relationships may play an important role in regional land surface parameterisation.

## Discussion

Approaches based on topographic predictors of LAI were examined in Chapter 4, and appear to offer stronger relationships than EO based approaches. We report at ~50% stronger relationship between LAI and topography than LAI and NDVI. In particular, scale dependent relationships between elevation and shelter (as quantified by TOPEX) are good predictors of LAI.

It is important to remember the dictum that correlation does not imply causation, and whilst statistical relationships between LAI and topography may be useful in a descriptive sense, it is important to develop a mechanistic understanding of the ecophysiology at work in order to successfully implement these findings in an operational sense. It is likely that the large-scale elevation trend is related to temperature, whilst the microscale relationship between LAI and exposure may be explained by local variations in snow accumulation, thaw dates and hence soil nutrient distributions (Wielgolaski et al., 1975). Further experimental work at Abisko is necessary to build process-based understanding of these issues for future modelling work.

We compared various methods of regionalising LAI in the Arctic tundra, to see if statistically optimal interpolation techniques such as Kriging could outperform simpler and computationally cheaper regression techniques. We report broadly similar interpolation skill for various Kriging techniques, inverse distance weighting (IDW) and linear regression, despite the utilisation of different combinations of data streams.

Despite results in chapter 3, where geostatistical methods provided no improvement in interpolation skill over simpler upscaling techniques, geostatistics are likely to remain an important part of any regionalised modelling activity. A common misconception about geostatistical methods is that they are limited to smoothly varying Gaussian fields with a constant mean: Geostatistical methods exist to deal with non-normality (Armstrong and Matheron, 1986a; Armstrong and Matheron, 1986b), discrete

## Discussion

boundary transitions (Goovaerts and Journel, 1995; Heuvelink and Webster, 2001) and secondary covariates (Goovaerts, 2000; Hudson and Wackernagel, 1994).

Geostatistics provide the ability to assimilate autocorrelated observations into arbitrarily complex regression models, making the techniques an ideal choice for future studies combining PFTs derived from EO vegetation classifications, radiance derived vegetation indices and topographic trends. Mixed effect type models (Pinheiro and Bates, 2000) are readily specified with slight modifications to the standard geostatistical methodologies (see chapter 7.8.2 and 7.9.1.i), providing a framework for the integration of such disparate data sources. Thus geostatistics remain a relevant research area given the significant advantage of offering spatial estimates of parameter variance, which is a key step to developing the potential for spatial data assimilation.

In the second section of the thesis, we turned our attention to problems of upscaling meteorological driving variables. There is a long tradition of such interpolation work in the literature (Ashraf et al., 1997; Hudson and Wackernagel, 1994; Running et al., 1987; Thiessen, 1911; Thornton et al., 1997), although in general this has been attempted at coarser spatio-temporal resolutions (Fuentes et al., 2006). Regionalising precipitation fields is considerably more problematic at fine temporal resolution (Thornton et al., 1997), and we report consistently poor regionalisation of rainfall fields (chapter 5). Geostatistics have previously been reported to provide slight improvements over other methods of regionalisation, and we find broadly similar results to those previously published (Diodato and Ceccarelli, 2005; Goovaerts, 2000): Although geostatistical methods provide an improved analysis over IDW approaches, the relative improvement may be small (Thornton et al., 1997); for example in the case of precipitation our geostatistical analysis provided worse results than IDW.

## Discussion

We implemented state of the art spatio-temporal geostatistical techniques to investigate potential improvements to high-resolution meteorological fields over simpler Kriging techniques that ignore temporal autocorrelation. In every case we report a poorer interpolation skill for the spatio-temporally explicit regionalisations (chapter 5). As such, one may question the merit of geostatistical techniques given their perceived complexity, computational cost and marginal improvement in interpolation skill. However, we reiterate the importance of attaching estimates of uncertainty to regionalised variables, which Kriging and its variants (geostatistical techniques subsume spline fitting methods (Matheron, 1981; Serra, 1987)) are uniquely able to achieve.

We examined the effect of *post hoc* temporal data aggregation on estimated fields of meteorology, and report a decrease in error as the size of temporal window increases. This led us to hypothesise that the somewhat large interpolation errors may cancel out over time in the model structure. Specifically, we hypothesised that integrating processes in the model structure would ‘smooth out’ errors, whilst rapid processes which react instantaneously to driving variables would display greater error (chapter 6). We report that temperature was the largest component of the meteorological uncertainty, supporting the hypothesis that instantaneous effects dominate the uncertainty of the NEE trajectory. Despite precipitation having the largest uncertainty and poorest  $r^2$  when compared with observations, the resultant impact on NEE uncertainty was minimal. We demonstrated the temporal buffering of uncertainty attributable to precipitation, and suggest this is because the effect of precipitation on vegetation is manifested through drought; an integrative effect related to soil texture and water holding capacity.

The second aim of the thesis was to compare the magnitude of the effects of parameterisation and driver uncertainty on the total C budget. In chapter 6 we present an analysis with significant novelty, being the first study to implement an Ensemble

## Discussion

Kalman Filter (ENKF) to provide constrained parameter uncertainty distributions for a C dynamics model. Furthermore it is the first study to implement spatio-temporal geostatistical simulation *via* the product-sum covariance model (De Cesare et al., 2001) to provide an ensemble of probable meteorological scenarios. We report that despite sizable uncertainties in driver fields, only small NEE uncertainties were attributable to meteorology. This was true even under extreme data scarcity, simulated by ignoring all available data <100 km from the study site. Conversely, the uncertainties associated with parameterisation accounted for up to 50% of the total NEE predicted by the model.

The ability of data assimilation to correct parameters based on incoming data streams was demonstrated in Chapter 6, indicating that the tools necessary to constrain and reduce the uncertainties associated with the exchange surface are already in place. The ability of DA techniques to reduce model uncertainty has been proven elsewhere (Evensen, 1994; Quaife et al., 2008; Williams et al., 2005), and such model correction is likely to continue to be important with the arrival of new EO data streams e.g. NASA's Orbiting Carbon Observatory (OCO) (Crisp and Johnson, 2005), and ESA's forthcoming Earth explorer mission (e.g. BIOMASS) (Bensi et al., 2007). It is the opinion of the author that geostatistical and DA technologies are amongst the most promising and relevant areas of C cycle science at present, and are worthy of considerable attention in the future.

## 8.1 References

- Armstrong, M. and Matheron, G., 1986a. Disjunctive kriging revisited 1. *Mathematical Geology*, 18(8): 711-728.
- Armstrong, M. and Matheron, G., 1986b. Disjunctive kriging revisited 2. *Mathematical Geology*, 18(8): 729-742.
- Ashraf, M., Loftis, J.C. and Hubbard, K.G., 1997. Application of geostatistics to evaluate partial weather station networks. *Agricultural and Forest Meteorology*, 84(3-4): 255-271.
- Bensi, P. et al., 2007. A new Earth explorer the third cycle of core Earth explorers. *ESA Bulletin*, 131: 31-36.
- Crisp, D. and Johnson, C., 2005. The orbiting carbon observatory mission. *Acta Astronautica*, 56: 193-197.
- De Cesare, L., Myers, D.E. and Posa, D., 2001. Estimating and modeling space-time correlation structures. *Statistics & Probability Letters*, 51(1): 9-14.
- Diodato, N. and Ceccarelli, M., 2005. Interpolation processes using multivariate geostatistics for mapping climatological precipitation mean in the Sannio mountains (southern Italy). *Earth Surface Processes and Landforms*, 30: 259-268.
- Evensen, G., 1994. Sequential data assimilation with a non-linear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *Journal of Geophysical Research*, 99: 10143-10162.
- Fuentes, M., Kittel, T.G.F. and Nychka, D., 2006. Sensitivity of ecological models to their climate drivers: Statistical ensembles for forcing. *Ecological Applications*, 16(1): 99-116.
- Goovaerts, P., 2000. Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall. *Journal of Hydrology*, 228(1-2): 113-129.
- Goovaerts, P. and Journel, A.G., 1995. Integrating soil map information in modeling the spatial variation of continuous soil properties. *European Journal of Soil Science*, 46(3): 397-414.
- Heuvelink, G.B.M. and Webster, R., 2001. Modelling soil variation: past, present, and future. *Geoderma*, 100(3-4): 269-301.
- Hudson, G. and Wackernagel, H., 1994. Mapping temperature using kriging with external drift - theory and an example from Scotland. *International Journal of Climatology*, 14(1): 77-91.
- Matheron, G., 1981. Splines and Kriging: Their Formal Equivalence. In: D.F. Merriam (Editor), *Down-to-Earth Statistics: Solutions Looking For Geological Problems: Syracuse University Geological Contributions*, pp. p.77-95.
- Pinheiro, J.C. and Bates, D.M., 2000. *Mixed-Effects Models in S and S-PLUS*. Statistics and computing. Springer, New York, xvi, 528 p. pp.
- Quaife, T. et al., 2008. Assimilating canopy reflectance data into an ecosystem model with an ensemble Kalman filter. *Remote Sensing of Environment*, 112: 1347-1364.
- Running, S.W., Nemani, R.R. and Hungerford, R.D., 1987. Extrapolation of synoptic meteorological data in mountainous terrain and Its use for simulating forest evapotranspiration and photosynthesis. *Canadian Journal of Forest Research-Revue Canadienne De Recherche Forestiere*, 17(6): 472-483.

## Discussion

- Serra, J., 1987. Comments on "Geostatistics: Models and tools for the Earth sciences" by A. G. *Journal. Mathematical Geology*, 19(4): 359-355.
- Thiessen, A.H., 1911. Precipitation averages for large areas. *Monthly Weather Review*, 39(7): 1082-1089.
- Thornton, P.E., Running, S.W. and White, M.A., 1997. Generating surfaces of daily meteorological variables over large regions of complex terrain. *Journal of Hydrology*, 190(3-4): 214-251.
- Wielgolaski, F.E., Kallio, P. and Rosswall, T., 1975. *Fennoscandian Tundra Ecosystems: Part 1 Plants and Microorganisms*. Ecological studies; v. 16-17, 1. Springer-Verlag, Berlin.
- Williams, M., Bell, R., Spadavecchia, L., Street, L.E. and van Wijk, M.T., 2008. Upscaling leaf area index in an Arctic landscape using multi-scale reflectance observations. *Global Change Biology*: DOI: 10.1111/j.1365-2486.2008.01590.x.
- Williams, M., Schwarz, P.A., Law, B.E., Irvine, J. and Kurpius, M.R., 2005. An improved analysis of forest carbon dynamics using data assimilation. *Global Change Biology*, 11(1): 89-105.
- Woodward, F.I., Smith, T.M. and Emanuel, W.R., 1995. A global primary productivity and phytogeography model. *Global Biogeochemical Cycles*, 9: 471-490.

## 9. Conclusions

The dominant source of uncertainty for regional scale C models is the parameterisation of the land surface. Critically, LAI and foliar N (which may be closely coupled to LAI) must be adequately specified at the working scale to produce reasonable estimates of NEE. EO data sources alone are inadequate to characterise the rapid transitions in LAI typical of Tundra ecosystems and apparent in high-resolution studies. The use of topographic indices derived from high quality DEMs such as those derived by LIDAR may go some way to improving land surface parameterisation. Despite issues of bias in meteorological upscaling, driver uncertainty contributes only marginally to the net uncertainty in C, even in cases of extreme data scarcity. We therefore conclude that future studies should concentrate resources on improving regionalisation of land surface parameters, although an analysis of driver uncertainty is advisable. State of the art spatio-temporal interpolation techniques did not improve driver surface accuracy over those that ignore temporal autocorrelation. We conclude that simpler solutions to upscaling are preferable in terms of computational cost and quality of output. Geostatistical techniques are essential for the calculation of surface error statistics, which are a key step towards regional scale DA implementation; we therefore suggest that regression based techniques or IDW are unsuitable for such studies. DA techniques have proven useful in the correction of model parameterisation, and are likely to provide improvements in model uncertainty in the future, especially if undertaken in 'online' mode such that assimilated observations can adjust parameter trajectories over time.