19th International Congress on Modelling and Simulation, Perth, Australia, 12–16 December 2011
http://mssanz.org.au/modsim2011

Determining the impact of reservoir water transfers on water quality using advanced methods

Roberta Fornarelli[a], **Jason P. Antenucci**[b] **and Clelia L. Marti**[a]

[a] *Centre for Water Research, The University of Western Australia, M023, 35 Stirling Hwy, Crawley 6009, Western Australia, Australia*

[b] *Hatch Associates, 144 Stirling St, Perth, Western Australia, Australia*

Email: fornarel@cwr.uwa.edu.au

Abstract: Inter-basin water transfers in reservoir systems, usually driven by purely economic purposes (e.g. water quantity objectives such as water supply and hydropower production), can have complex physical, chemical, hydrological and biological implications both in the upstream and downstream reservoir. Phytoplankton growth is a complex process, usually driven by internal dynamics and mainly dependent on nutrient concentrations, temperature and light availability. When water is transferred from an upstream to a downstream reservoir, phytoplankton cells can potentially be transported, thus causing changes in the phytoplankton community. Determining if a causal relationship exists between the phytoplankton flux from the upstream reservoir and the measured phytoplankton biovolume in the downstream reservoir becomes an important step towards the understanding and management of reservoir systems.

In the present study, simple and novel statistical methods (e.g. randomized intervention analysis and Granger causality), based on time series analysis and linear regression, were used to detect trends of phytoplankton biovolume and causal relationships between two interconnected reservoirs, thus providing significant evidence of how phytoplankton biovolume was affected by the magnitude of water transfers. Our analysis was supported by a weekly dataset of phytoplankton biovolume in both the reservoirs and by daily water quantity data (transferred flow rate from the upstream to the downstream reservoir). The studied scheme is the Shoalhaven System, Australia, which was built in the 1970s as a water supply and hydropower generation system. All the applied methods are suitable to be used in any kind of connected water system, provided a long-term dataset is available.

Two different time periods, characterized by low and high water transfers between the reservoirs, were identified and a Granger causality statistical test was applied on each of them. A causal relationship was found only during the high transfer period, i.e. when the transferred flow rate exceeded a certain threshold. This result demonstrates that the phytoplankton flux due to water transfers was one of the causes of the observed phytoplankton biovolume in the downstream reservoir. Therefore the increase of the observed biovolume in the downstream reservoir during the high transfer period was due to a combination of processes, i.e. the transport of phytoplankton cells from the upstream reservoir, a seeding effect of the imported phytoplankton and internal processes. The same procedure was applied on specific phytoplankton groups, i.e. diatoms, chlorophytes and cyanobacteria, testing the hypothesis of a causal relationship between measured biovolume in the downstream reservoir and flux from the upstream reservoir via water transfers. It was demonstrated that diatoms was the only phytoplankton group experiencing a causal relationship with the diatom flux from the upstream reservoir.

The application of the Granger causality test to ecological time series in the context of reservoir systems revealed to be a valuable statistical contribution to the understanding and management of connected water bodies. Future analyses will concentrate on broader applications of this methodology in the context of the understanding and management of reservoir systems.

Keywords: *water transfers, water quality, phytoplankton biovolume, Granger causality*

1. INTRODUCTION

Inter-basin water transfers, usually managed according to water quantity purposes (e.g. drinking water supply and hydropower production), can have complex biological implications both in the upstream and downstream reservoir mainly due to physical and biological differences between the connected systems and to the magnitude, frequency and duration of transfers (Soulsby et al., 1999, Gibbins et al., 2000).

Phytoplankton biovolume is a fundamental water quality variable in lakes and reservoirs, representing the ecosystem's status in terms of eutrophication trends and algal blooms. Phytoplankton biovolume is directly connected to phytoplankton growth, thus depends on internal processes involving nutrient concentrations, water temperature and light availability (Hypsey et al., 2007). However, in a reservoir system, the management of water transfers between two reservoirs could potentially affect the water quality, and thus also the phytoplankton biovolume, in the downstream reservoir. Within this context, determining the presence of a causal relationship between the water quality in the two reservoirs is a fundamental step towards a comprehensive understanding and management of the reservoir system.

Traditional regression and correlation analyses are not sufficient in determining a causal relationship between two variables (Kaufmann and Stern 1997, Wang et al., 2004). Instead, the Granger causality statistical test (Granger, 1969) is a well-established methodology to estimate the causal dependence between two time series. Granger causality tests are based on the notion of predictability, in particular "*a variable X is causal for another variable Y if knowledge of the past history of X is useful for predicting the future state of Y over and above knowledge of the past history of Y itself*" (Mosedale et al., 2006). Following this definition, the variable Y is said to be Granger caused by X if its prediction improves by including past values of X as a predictor. This method has been widely used in economic studies and, in the last twenty years, a few applications have also been made in environmental context for climate change studies (Kaufmann and Stern 1997, Salvucci et al., 2002, Wang et al., 2006, Mosedale et al., 2006, Elsner, 2007). The major limitation of this methodology is that the detection of Granger causality does not automatically involve the presence of a physical causal mechanism, and also the obtained results strictly depend on which conditioning variables are used (Kaufmann and Stern 1997). Therefore, caution is required in interpreting the statistical results in a physically meaningful way (Wang et al., 2006).

The purpose of this paper is to test the presence, or not, of a causal relationship between phytoplankton biovolume measured in two reservoirs, connected via inter-basin water transfers. The magnitude of water transfers changed throughout the study period, thus two different periods characterized by low and high water transfers were identified. The Granger causality statistical test was applied to each of these periods, using a dataset of weekly phytoplankton biovolume. To the best of our knowledge, the present study represents the first application of Granger causality to biological time series of phytoplankton data. Given the existing physical connection between the two reservoirs, Granger causality seems to be a promising tool and a valuable statistical contribution to the understanding and management of connected water bodies such as reservoir systems.

2. METHODS

2.1. Description of the system and data availability

Fitzroy Falls Reservoir (Fig. 1) is a small, shallow reservoir, part of the Shoalhaven System (New South Wales, Australia). It is artificially connected to Lake Yarrunga (Fig. 1), from which it receives pumped water through inter-basin water transfers for water supply and hydropower generation. Bendeela pondage is a small reservoir, sited between Lake Yarrunga and Fitzroy Falls to aid hydropower generation. Its effect on the water quality of Fitzroy Falls was neglected in this study because of its low retention time (2 days). The transferred inflow from Lake Yarrunga is Fitzroy Falls' biggest inflow; the variability of the magnitude and frequency of the transferred inflow is high and is entirely regulated by the management of the transfer scheme between Lake Yarrunga and Fitzroy Falls. Note that, in the following analysis, Fitzroy Falls was referred to as the "downstream reservoir" and Lake Yarrunga as the "upstream reservoir".

Daily inflows and outflows in Fitzroy Falls were available from January 2001 to February 2010. Surface phytoplankton biovolume data were available in one location in Fitzroy Falls close to the dam wall (Fig. 1) for the same period. Data were available at genus level and total phytoplankton biovolume was obtained by summing up the biovolume of the measured genera. The same dataset was measured for the same period in Lake Yarrunga at a monitoring station 1 km downstream of the pumping station (Fig. 1), and was used to represent the phytoplankton biovolume transported in Fitzroy Falls by the water transfers. Note that phytoplankton biovolume in the two reservoirs was measured weekly for each year except during austral winter months (from June to August) when it was measured once a month. Therefore a constant value, equal to the monthly measurement available, was considered across each week of the corresponding month. Phytoplankton fluxes towards Fitzroy Falls were calculated by multiplying the weekly averaged transferred flow rate by the weekly phytoplankton biovolume measured in Lake Yarrunga. The entire dataset was provided by the Sydney Catchment Authority and all measurements were conducted using APHA standard protocols (American Public Health Association, 2005).

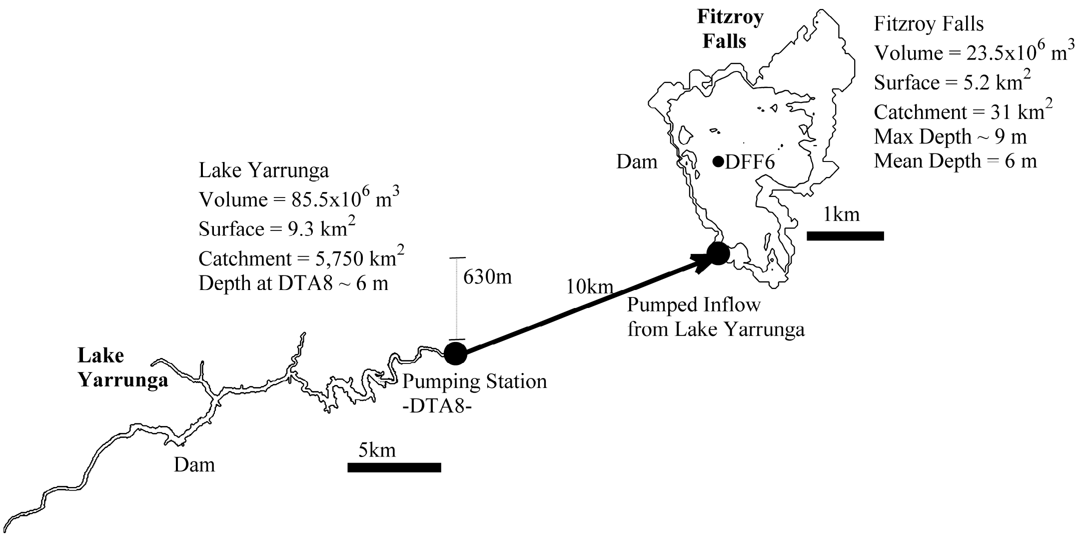

Figure 1. Lake Yarrunga and Fitzroy Falls Reservoir connection and physical characteristics. Phytoplankton data were available at two monitoring stations (one in Lake Yarrunga, DTA8, and one in Fitzroy Falls, DFF6). Fitzroy Falls contours at 0, 4, 8 m depth. Elevation difference: 630 m; distance between the two reservoirs: 10 km. Figure not to scale.

2.2. Water transfers management

The magnitude of water transfers from Lake Yarrunga to Fitzroy Falls strongly varied in the last ten years, potentially affecting the water quality in Fitzroy Falls. Relatively low water transfers occurred between the two reservoirs from June 2001 to May 2003; the average daily inflow from Lake Yarrunga to Fitzroy Falls was $100 \pm 125 \times 10^3$ m^{3}d^{-1} (mean ± std, 200 days retention time). These years were referred to as the “low transfer period”. In turn, the “high transfer period” was defined from June 2006 to May 2008, with an average daily inflow of $545 \pm 329 \times 10^3$ m^{3}d^{-1} (mean ± std, 37 days retention time). Water retention time in Fitzroy Falls was calculated as the ratio between the storage volume and the total inflow, both averaged over the entire low (or high) transfer period.

2.3. Randomized Intervention Analysis

The effect of water transfers on the phytoplankton biovolume in Fitzroy Falls was evaluated through a direct comparison between the phytoplankton biovolume measured during low and high transfer periods. To determine the statistical significance of the difference in the mean concentrations, Randomized Intervention Analysis (RIA) was used (Carpenter et al., 1989). The actual difference between mean concentrations during high and low transfer periods was ranked along with its probability distribution to produce a probability value (p-value). The probability distribution was obtained by random permutations of biovolume time series during the two time periods.

2.4. Granger Causality

The Granger causality approach was used to test the hypothesis that the phytoplankton flux transferred from the upstream reservoir does not cause the phytoplankton biovolume measured in the downstream one (null hypothesis). The procedure suggested by Kaufmann and Stern (1997) and Wang et al. (2006) was followed and applied to the weekly biovolume time series during the low and high transfer period.

In the first step, the following statistical model (Eq. 1) was defined:

$$BV_{DW}(t) = a_1 t + \sum_{i=1}^{p} b_i BV_{DW}(t-i) + \sum_{j=0}^{p} c_j Flux_{UP}(t-j) + \varepsilon_t \tag{1}$$

where BV_{DW} is the biovolume in the downstream reservoir, $Flux_{UP}$ is the transferred phytoplankton flux from the upstream reservoir, ε_t is a normally distributed random error term, p is the model order, a_1, b_i and c_j are regression coefficients, and t is the time interval (week). The time component, a_1t, was added as a linear filter to correct the non-stationary behavior of the time series (Kaufmann and Stern, 1997, Shumway, 1988). Equation 1 represents the unrestricted model, where the biovolume in the downstream reservoir depends on itself as well as on the phytoplankton flux from the upstream reservoir (i.e. assuming the time series were Granger causal).

Equation 2 constitutes the restricted model:

$$BV_{DW}(t) = d_1 t + \sum_{i=1}^{p} f_i BV_{DW}(t-i) + \eta_t \tag{2}$$

where now η_t is a normally distributed random error term, d_1 and f_j are regression coefficients. The dependence on the biovolume measured in the upstream reservoir was eliminated by forcing the regression coefficients c_j to zero, thus considering that the biovolume of the downstream reservoir depended only on previous values of itself (i.e. assuming the series were not Granger causal).

Both models (Eqs. 1 and 2) were solved by ordinary least squares regression (Wang et al., 2006) and the best model order p was chosen according to the minimization of Akaike's Information Criterion and Bayesian Information Criterion (Shumway, 1988).

The significance of the restriction applied to model 2 was tested applying a F-test based on the following statistic W (Eq. 3, Kaufmann and Stern, 1997):

$$W = \frac{(RSSr - RSSu)/s}{(RSSu)/(T-k)} \tag{3}$$

where T is the number of observations, k is the number of parameters in the unrestricted model (Eq. 1), s is the number of parameters that are restricted to zero in the restricted model (c_j), RSS is the sum of squared residuals from restricted ($RSSr$) and unrestricted ($RSSu$) models. The test statistic was compared with the tabulated value of the F distribution (F_{critic}), at 5% significance level, with s and T-k degrees of freedom; if $W > F_{critic}$, the hypothesis was rejected with $p < 0.05$ and thus the phytoplankton flux from the upstream reservoir was said to Granger cause the biovolume in the downstream one.

3. RESULTS AND DISCUSSION

Phytoplankton biovolume in Fitzroy Falls significantly increased (p-value = 0 from RIA analysis, Carpenter et al., 1989) from an average of 1.66 ± 0.8 mm^3L^{-1} during low transfer period to 3.14 ± 1.7 mm^3L^{-1} during the high transfer period. The observed change in the phytoplankton biovolume could be due to several factors, e.g., dissimilar nutrient concentrations, difference in temperature, stratification regime and retention time between the two periods. However, in a system of interconnected reservoirs, the seeding effect and the direct transfer of phytoplankton cells from the upstream to the downstream reservoir can play a fundamental role in determining the observed concentrations in the downstream reservoir. Within this context, the application of a Granger causality test during low and high transfer periods might be helpful in clarifying the relationship between observed biovolume in the downstream reservoir and phytoplankton flux transported from the upstream one.

The hypothesis to be tested was that the phytoplankton flux from the upstream reservoir does not cause the measured biovolume in the downstream one. Weekly time series of phytoplankton biovolume and phytoplankton flux were used both for low and high transfer period, resulting in a total of 104 observations

(T) for each period. Data were fitted following the unrestricted and restricted models (Eqs. 1 and 2, respectively). According to the minimization of Akaike's Information Criterion and Bayesian Information Criterion (Shumway, 1988), the model order 2 was chosen during the low transfer period and model order 1 during the high transfer period. The results of the regressions are shown in Figure 2a and b.

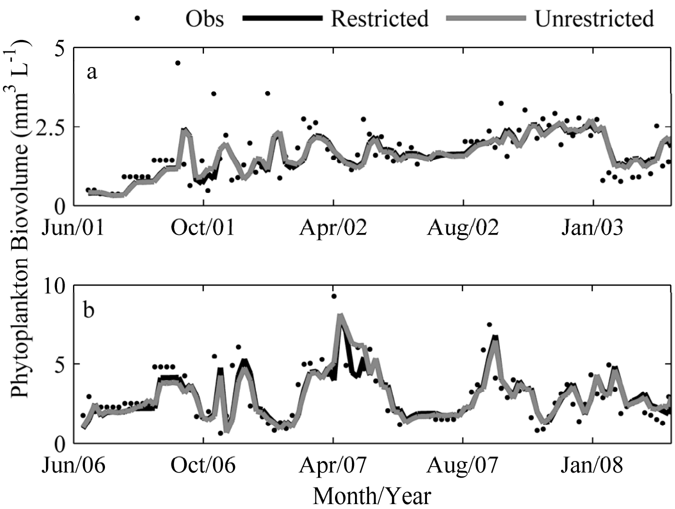

Figure 2. Phytoplankton time series measured in the downstream reservoir during low (a) and high (b) transfer period (dot points). For each period, simulated trajectories of restricted (black line) and unrestricted (grey line) regression models are shown.

During the low transfer period, the performance indices, i.e. R^2 and index of agreement D (Willmott, 1982), of the restricted and unrestricted models (Fig. 3) were similar; both the restricted and unrestricted model captured about 14% of biovolume variance with an index of agreement of about 0.70. On the contrary, during the high transfer period the performance indices (Fig. 3) differed between the restricted and unrestricted model, with the latter providing a better fit (R^2 of about 47%, D about 0.83).

The difference of R^2 between the unrestricted and restricted models (Fig. 3) represents the explanatory power exclusively provided by the phytoplankton flux (Wang et al., 2006). During the low transfer period, the phytoplankton flux alone accounted for less than 1% of the variance (Fig. 3a) while, during the high transfer period, its explanatory power increased to 4.3% (Fig. 3a). To statistically test the different performances of the restricted versus the unrestricted model in the two transfer periods, an F-test was applied following Eq. 3 (Table 1). The null hypothesis is that the phytoplankton flux transferred from the upstream reservoir does not cause the phytoplankton biovolume measured in the downstream one, therefore the performances of the unrestricted and restricted models are statistically equal.

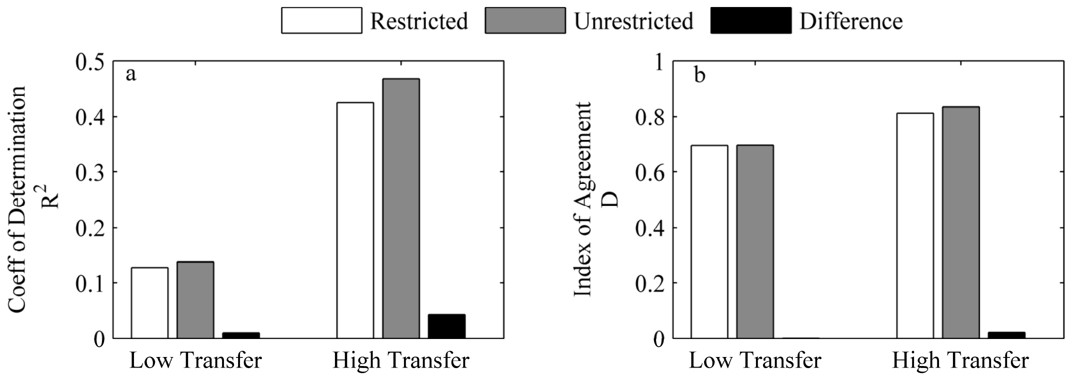

Figure 3. Model performance in term a) of R^2 and b) D of the restricted and unrestricted models during the low and high transfer periods. The black bars represent the difference of R^2 and D between unrestricted and restricted models.

During the low transfer period, the number of parameters in the unrestricted model (k) was 6 (refer to Eq. 1 with model order 2), thus the number of parameters restricted to zero (s) was 3. Therefore the test statistic W was compared with the F distribution with 3 and 98 degrees of freedom: the hypothesis was accepted (Table 1) as $W < F_{critic}$, thus the phytoplankton flux from the upstream reservoir did not cause the phytoplankton biovolume in the downstream reservoir. During the high transfer period, the test statistic was compared with the F distribution with 2 and 100 degrees of freedom: the hypothesis was rejected (Table 1) and the phytoplankton flux from the upstream reservoir did cause the phytoplankton biovolume in the downstream reservoir.

Table 1. Results of Granger causality test during low and high transfer period. Values in bold indicate the test was rejected indicating that the causal relationship is present. W = test statistic; v1 and v2 = degrees of freedom of F distribution; F_{critic} = critical value of F distribution at 5% significance level.

	W	v1	v2	*Fcritic*
Low Transfer Period	0.381	3	98	2.697
High Transfer Period	**4.045**	**2**	**100**	**3.087**

The results of Granger causality test demonstrated that, when water transfers were below a certain value of flow rate, the phytoplankton biovolume in the downstream reservoir was not seeded by the phytoplankton flux coming from the water transfers. Internal processes and other forcing factors contributed to the observed biovolume. On the other hand, when water transfers increased, the contribution of the phytoplankton biovolume transported from the upstream reservoir was causing the observed value in the downstream reservoir.

This result provides an important outlook into the understanding and management of the system, representing the first step towards more detailed analyses on the phytoplankton community. The procedure we applied so far can be extended in order to understand which phytoplankton group was responsible for the observed causal relationship. As an example, we applied the Granger causality test on the time series of diatom, chlorophyte and cyanobacteria biovolume. The null hypothesis tested that the diatom (or chlorophyte or cyanobacteria) flux transferred from the upstream reservoir does not cause the diatom (or chlorophyte or cyanobacteria) biovolume measured in the downstream one. The final results are shown in Table 2.

Table 2. Results of the Granger causality test during the low and high transfer periods applied to diatom, cyanobacteria and chlorophyte biovolume.

	Low Transfer Period	High Transfer Period
Diatoms biovolume	Non-causal	**Causal**
Chlorophyte biovolume	Non-causal	Non-causal
Cyanobacteria biovolume	Non-causal	Non-causal

During the low transfer period, none of the three phytoplankton groups was caused by the flux of biovolume transferred from the upstream reservoir (Table 2), confirming the results obtained for the total phytoplankton biovolume. Whenever the magnitude of water transfers was below a certain value, internal processes rather than cells transport from the upstream reservoir were responsible for the observed biovolume. On the contrary, during high transfer period, chlorophyte and cyanobacteria biovolume were still unrelated to the flux from the upstream reservoir while diatom biovolume was Granger caused by the flux of diatoms. High water transfers increased the flux of phytoplankton, i.e. diatoms, from the upstream reservoirs, thus contributing through direct transport of phytoplankton cells to the biovolume measured in the downstream reservoir. This analysis has potentially important implications in the management of water transfers: it demonstrated that water transfers were responsible for part of the measured phytoplankton biovolume, exerting their control only on the diatom community without affecting cyanobacteria and chlorophytes via cells transport.

The application of Granger causality on this specific case study was successful in detecting trends and is a promising tool for future application in such systems. Moreover, one of the main limitations of Granger causality (difficult interpretation of the statistical results in a physically meaningful way; Kaufmann and Stern, 1997, Wang et al., 2006) can be overcome by applying it in interconnected reservoir systems where the transport of substances, thus causal relationships, from one reservoir to another can naturally happen.

4. CONCLUSIONS

The present study applied the Granger causality test in order to explain the observed phytoplankton biovolume in the receiving end of two interconnected reservoirs. The presence of a Granger causal relationship between phytoplankton biovolume in the downstream reservoir and the flux of phytoplankton transferred from the upstream reservoir was found only when high water transfers occurred between the reservoirs. This causal relationship was ascribed to the transport of diatoms cells. The results of this study lead to the following conclusions: i) the connection between reservoirs through water transfers can potentially affect the water quality of the downstream reservoir by simple transport of phytoplankton cells; when the water transfers increased, the water quality as observed in the downstream reservoir was affected by both internal processes and transport of material from the upstream reservoir; and ii) the Granger causality test revealed to be an efficient method in analyzing the connection between reservoirs and can be easily applicable to any kind of reservoir-reservoir or river-reservoir system, provided that relevant datasets are available in both systems. These results have important implications in the understanding and management of reservoir systems: the presence or absence of causal relationships between the water quality of interconnected reservoirs can lead to specific decisions about the water transfer management strategies (e.g. the timing and quantity of transfers).

ACKNOWLEDGMENTS

The first author was the recipient of a Scholarship for International Research Fees from the University of Western Australia and of a University International Stipend from the Centre for Water Research. We gratefully acknowledge the support of the Sydney Catchment Authority for providing access to the data. The findings, opinions and conclusions expressed herein are those of the authors and do not represent the views or opinions of the Sydney Catchment Authority or any person employed in the Sydney Catchment Authority Division. This article represents Centre for Water Research reference 2484-RF.

REFERENCES

American Public Health Association (2005). Standard Methods for the Examination of Water and Wastewater, 21st edn. American Public Health Association (APHA), American Water Works Association (AWWA) & Water Environment Federation (WEF).

Carpenter, S.R., Frost, T.M., Heisey, D., Kratz, T.K., (1989). Randomized intervention analysis and the interpretation of whole-ecosystem experiments. *Ecology*, 70(4), 1142-1152.

Elsner, J.B., (2007). Granger causality and Atlantic hurricanes. Tellus, 59A, 476-485.

Gibbins, C.N., Jeffries, M.J., and Soulsby, C., (2000). Impacts of an inter-basin water transfer: distribution and abundance of Micronecta poweri (Insecta: Corixidae) in the River Wear, north-east England. *Aquatic Conservation: Marine and Freshwater Ecosystem*, 10, 103-115.

Granger, C. W. J., (1969). Investigating causal relations by econometric models and cross spectral models. *Econometrica*, 37, 424–438.

Hipsey, M.R., Romero, J.R., Antenucci, J.P., and Imberger, J., (2007). The Computational Aquatic Ecosystem Dynamics Model (CAEDYM): a versatile water quality model for coupling with hydrodynamic drivers. *Proceedings of the 7th International Conference on Hydroinformatics*, Nice, France, 03-08 September 2006, Vol 1., 526-533.

Kaufmann, R.K., and Stern, D.I., (1997). Evidence for human influence on climate from hemispheric temperature relations. *Nature*, 388, 39-44.

Mosedale, T.J., and Stephenson, D.B., (2006). Granger causality of coupled climate processes: ocean feedback on the north Atlantic oscillation. *Journal of Climate*, 19, 1182-1194.

Salvucci, G.D., Saleem, J.A., and Kaufmann, R., (2002). Investigated soil moisture feedbacks on precipitation with tests of Granger causality. *Advances in Water Resources*, 25, 1305-1312.

Shumway, R. H., (1988). Applied statistical time series analysis, Prentice-Hall International Editions, Inc. A Division of Simon & Schuster Englewood Cliffs, NJ 07632

Soulsby, C., Gibbins, C.N., and Robins, T., (1999). Inter-basin water transfers and drought management in the Kielder/Derwent system. *Journal of the Chartered Institution of Water and Environmental Management*, 13, 213-223.

Wang, W., Anderson, B.T., Kaufmann, R.K., and Myneni, R.B., (2004). The relation between the North atlantic oscillation and SST's in the North Atlantic basin. *Journal of Climate*, 17, 4752-4759.

Wang, W., Anderson, B.T., Phillips, N., and Kaufmann, R.K., (2006). Feedbacks of vegetation on summertime climate variability over the North American grasslands. Part I: statistical analysis. *Earth Interactions*, 10, 17.

Willmott, C.J., (1982). Some comments on the evaluation of model performance. *Bulletin of American Meteorological Society*, 63, 1309-1313.