Is Victoria's riverine salinity resilient to extreme hydrological events?

N. Kho, T. Peterson & A. Lintern
Department of Civil Engineering, Monash University, Australia

ABSTRACT: Catchment hydrologic responses have been widely assumed to always recover from disturbances such as extreme drought. It is uncertain if this holds true for stream water quality, which is evaluated by indicators such as salinity. To investigate this question, Hidden Markov Models (HMMs), accounting for rainfall as a predictor, were used to analyse monthly riverine salinity fluxes at 8 sites in Victoria, to understand the impact of the Millennium Drought on salinity states. Two-state models, where salinity loads varied between "normal" and "low" states, were found to better predict in-stream salinity compared to single-state models. The model results showed that changes to a low salinity state lagged behind changes to a low runoff state. As groundwater is the main source of salinity in these catchments, this suggests that reductions in groundwater flows into rivers occur as a result of the shift to a lower runoff state. These findings highlight the limited resilience of a catchment's salinity to drought. It is possible that other water quality constituents will also show limited resilience. To improve model accuracy, more complex HMMs involving multivariate analysis are promising for future research. Understanding how easily water quality in catchments shifts to different states due to extreme hydrologic events allows proper catchment management, thus preventing water quality deterioration from causing ecological harm.

1 INTRODUCTION

Good quality water is essential for all life. Otherwise, we see disease in communities lacking safe drinking water, and degradation of polluted ecosystems. For instance, some freshwater organisms cannot tolerate the increasing salinity observed in certain rivers, harming biodiversity (Cañedo-Argüelles et al., 2013). Highly saline rivers unsuitable for irrigation or drinking purposes may also cause further socio-economic damage. Proper management of water resources mitigates such effects, but this requires understanding the patterns and processes that cause water quality to vary spatially and temporally. In understanding this, previous studies have shown that water quality processes and average constituent levels are driven by catchment characteristics (Lintern et al., 2018a; Musolff et al., 2015).

Variations in water quality over time have also been linked to variations in water quantity. The relationship between in-stream constituent concentration and streamflow i.e., the C-Q relationship, has been previously studied (Bieroza et al., 2018; Moatar et al., 2017; Musolff et al., 2015). This C-Q relationship is not always static, meaning that constituent concentrations for a given streamflow may at times follow a different trend. This has been shown to occur during short-term events, such as floods (Rue et al., 2017) and storms (Knapp et al., 2020; Minaudo et al., 2019, where the C-Q relationship itself was found to differ from average conditions. During high-flow events, for nutrient-based constituents, the slope of the C-Q relationship may change in both directions, representing changes in mobilisation or dilution behaviour (Knapp et al., 2020; Minaudo et al., 2019). On the other hand, limited variability was observed for groundwater-based constituents during similar events (Knapp et al., 2020; Rue et al., 2017).

Short-term events, however, are not the only incidents experienced by catchments. Variations in long-term decadal scale disturbances cause changes in the typical patterns and amounts of water received, stored, and released, thus affecting all catchment processes. The Millennium Drought in south-eastern Australia (~1995-2009) was a clear example of a long-term, prolonged, and extreme hydrological event (van Dijk et al., 2013). During this period, catchments in these locations saw significant changes to their typical rainfall-runoff response, characterised by less runoff for any given rainfall than expected historically (Saft et al., 2015). A further pressing finding to this, was that even after the Millennium Drought had ended, this rainfall-runoff response change had not recovered in Victorian catchments, and was unlikely to recover soon (Peterson et al., in review). This suggests that when catchments experience a disturbance (such as a major hydrological event), hydrological processes may not necessarily recover (Peterson and Western, 2014). In this study, we call this phenomenon 'finite resilience'.

Holling (1973) first introduced the concept of ecosystem resilience, where for a system undergoing change, having some capacity to absorb these disturbances enables retainment of the same function and identity (Walker et al., 2004). In many natural systems with limited resilience, perturbations cause a dramatic change from one steady state system to another, with no change back in state, until some other significant forcing is able to cause such a reversion (Scheffer and Carpenter, 2003). Non-reverting changes support arguments that catchments can have more than one steady state, as predicted for vegetation-soil-moisture-groundwater storage interactions (Peterson et al., 2009, 2012; Peterson and Western, 2014).

Whilst resilience of water quantity in catchments has been previously investigated, we still have limited understanding of the impact of extreme hydrological events on water quality, including the impact of the Millennium Drought on Victorian stream water quality. By investigating water quality for similar state changes, we stand to gain a better understanding of how such extreme events impact water quality, as well as when and how recovery in water quality may occur. Furthermore, through the perspective of the mechanisms influencing stream water quality, such analysis could aid explanation of why catchment runoff behaviour has not recovered since the Millennium Drought.

In traditional analyses of water quality trends using C-Q methods, delineation of the periods of study prior to regression line analysis of the different periods is required. At some Victorian sites, previous work by Kho et al. (2020) in terms of riverine salt concentrations highlighted that there were statistically significant downwards shifts in the intercepts of electrical conductivity-Q trendlines from the pre-drought to post-drought phase. However, such methods, by relying on first pre-determining study time periods, prevent identification of more nuanced regime changes. Thus, there first exists a need for robust methods of change analysis of long-term water quality trends, particularly changes in the C-Q relationship in time and evolution of seasonality (Hirsch et al., 2010). Secondly, whilst being able to identify changes in streamflow water quality regimes in the past is important, by extension, understanding the likelihood and persistence of such changes is also important.

Hidden Markov Models (HMMs) can be used to explore these questions of state change. HMMs can detect hidden different water quality 'states' based on the data's distribution, and are highly flexible, allowing multiple parameters to be jointly analysed (Zucchini and MacDonald, 2009). They address the first need above by providing a robust method of change analysis. Furthermore, they provide an understanding of likelihoods of these changes, in contrast to other studies of water quality time series, such a Mann-Kendall trend and breakpoint analysis (Shafiei and McLoughlin, 2017). Whilst such studies were able to identify seasonal variability of nutrient loads, they were not able to directly address the probabilities and persistence of changes.

'Hidden' in the term 'Hidden Markov Model' suggests that only the outcomes from being in a state can be observed, but not the state itself. For example, during a state of drought, large rainfall events could still be observed, though this is less likely than during a non-drought period. Similarly, whilst we cannot determine directly what runoff state a catchment is in, we can use observed streamflow data to infer the most likely state (Peterson et al., in review). To predict these states, mixture models are fitted to the observations, such as by using Maximum Likelihood Estimation (Zucchini and MacDonald, 2009), and thus, observations coming from different states should come from different distributions that make up the mixture model. Hence, for each observation, there is a probability of it being in each state, i.e. the 'emission' probabilities. Furthermore, for each state, there is the possibility of remaining in this state, or switching to a new state, which are the 'transition' probabilities. These probabilities are then considered by a likelihood function in the HMMs. Finally, the most probable sequence of states based on these probabilities can be identified, typically by using the Viterbi Algorithm (Viterbi, 1967).

Within the field of hydrology, HMM variations and hybrids have been used in water quality studies undertaken in urban (Suchetana et al., 2019), marine (Rousseeuw et al., 2015), and riverine environments (Spezia et al., 2011) to identify changing water quality regimes. Spezia et al. (2011) showed that HMMs could be used to identify long-term changes in nitrogen species in Scottish river systems from time series data, though this analysis did not consider natural disturbances in terms of extreme climatic events. In terms of water quantity, Peterson et. al (in review) showed that runoff in south-east Australia, even after considering the impacts of precipitation, experienced long term declines due to drought which did not recover, highlighting that catchments may have limited resilience. This study brings together these two research themes to investigate if Victorian water quality has multiple steady states even after accounting for covariates, and if such states show recovery after an extreme hydrological event, namely, the Millennium Drought. Salinity in terms of total dissolved solids (TDS) was investigated, given the availability of this data in the region, and its ability to cause significant stress in freshwater systems if unmanaged (Cañedo-Argüelles et al. 2013).

1.1 *Hypothesis*

The null hypothesis of this study is that riverine salinity has only one state. To falsify this, this study investigates for the existence of multiple steady states in salinity, after accounting for independent drivers, namely precipitation. For this hypothesis to remain true, no evidence or conflicting evidence of multiple steady states would be expected from the results of this study. In such a case, catchments could be considered to always recover from extreme disturbances. However, should evidence of multiple steady states be obtained instead, this would indicate that the processes governing in-stream salinity could have limited resilience.

1.2 *Objective*

This aim of this study is to quantitatively assess if non-urban riverine salinity displays multiple steady states due to drought, as shown in streamflow (Peterson et al., in review). Should changes persist after drought within the new state, as opposed to recovery to the initial state post-event, the system can be said to display a finite resilience. As a result of this analysis, we will gain a better understanding of how stream salinity is likely to shift during extended droughts. This will also provide insight into the impact of rainfall, surface runoff and ground storage on temporal variability in river salinity. This study addresses the following three questions: (1) Is observed monthly stream salinity, once accounting for rainfall variability, best explained by single-state models? (2) How heavily is model performance affected by discrepancies in input data? (3) Given any changes in state, did salinity show evidence of non-recovery from prolonged drought?

1.3 *Scope*

In this study, eight non-urban Victorian sites were investigated through HMM analysis of monthly TDS loads, predicted by total monthly rainfall. This study only considers statistical models with the existence of two possible steady states, such as ‘High’ and ‘Normal’ salinity states, or ‘Normal’ and ‘Low’ states, relative to the reference year chosen to be ‘Normal’. Models with more than 2 states were not considered. This is as considering the existence of just two states is sufficient to identify whether there was a significant and sustained shift in state away from ‘typical’ conditions. To determine the total number of distinct riverine salinity steady states occurring in nature, is a second order-question, beyond the scope of this study.

2 MATERIALS AND METHODS

2.1 *Data*

2.1.1 *Data Format*

In-stream salinity is measured in terms of electrical conductivity (EC). A higher electrical conductivity is due to a higher concentration of aqueous ions available to conduct charges, reflecting higher in-stream salinity. Electrical conductivity data (μcm/s) obtained at frequent intervals (minimum time between readings of 2 hours) and streamflow data (ML/day) were downloaded from the WMIS website (State of Victoria DELWP, 2020). Catchment weighted rainfall data was obtained from area weighted grid precipitation data using the AWAPer R package (Peterson et al., 2020).

Water quality data for many constituents, including measures of EC, have traditionally been available through the testing of samples manually collected at regular time intervals. In Victoria, such data is commonly available ‘infrequently’ at monthly intervals across most monitored sites. More ‘frequent’ EC data is available at select sites across Victoria, where such EC measurements are obtained at a minimum of 2-hour intervals using automatic sampling. Consequently, this ‘frequent’ data is better able to capture the natural variability expected in riverine salinity.

Such continuous frequent data also allows for the estimation of salinity loads in terms of TDS for a site. In converting EC data (units of μcm/s) to total dissolved solid concentrations (units of mg/L), a multiplication factor of 0.6 can be used (Agriculture Victoria, 2018). Multiplication of TDS values by the instantaneous flow observed produced the instantaneous salt flux. Instantaneous flow data was readily available at frequent intervals, where frequent EC was also available. Such values of streamflow were linearly interpolated between records where no streamflow readings were coincident with EC readings, provided there had been at least one streamflow reading on that day. Instantaneous salt fluxes were then integrated over each month by the trapezoidal method (Fisher et al., 2016), giving the final monthly salinity loads.

In the HMM analysis, monthly TDS loads were predicted using total monthly rainfall. We chose not to use streamflow as a predictor because streamflow was already used to calculate the monthly TDS load. As a result, there is a spurious strong correlation between streamflow and TDS loads. Whilst runoff is consequential from rainfall, the rainfall-runoff relationship is not directly definitive, providing some degree of separation, thus likely minimising the introduction of such incorrect, pre-established correlation. In developing the HMM structure, the relationships between the natural log of both instantaneous EC concentration-streamflow (C-Q, Fig. 1) and monthly TDS flux-precipitation (flux-P, Fig. 2) were visually examined and seen to both be somewhat linear, with R^2 values of 0.36 and 0.46, respectively. It is noted that a slight heteroscedasticity is observed in the monthly TDS flux-precipitation model (Fig. 2), however the HMM structure assumes homoscedasticity, that is a constant variance. Nonetheless, by investigating

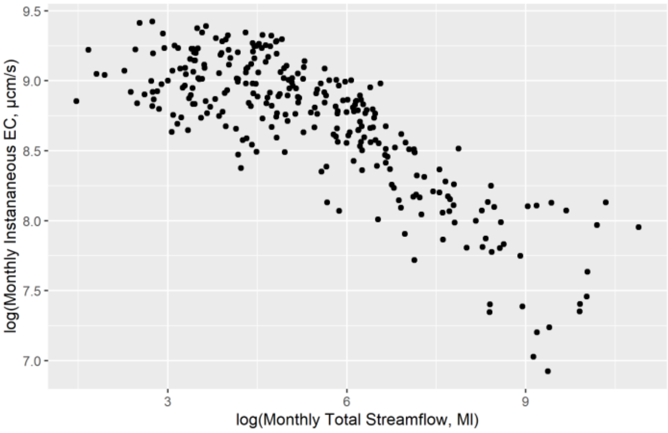

Figure 1. Monthly salinity measured as instantaneous point samples of electrical conductivity representative of the whole month vs. total streamflow for that month at Site 234201.

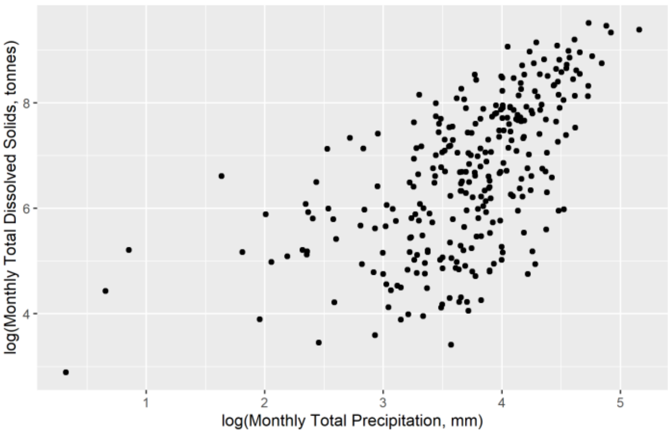

Figure 2. Monthly salinity measured as total dissolved solids vs. total precipitation for that month at Site 234201.

flux-P instead of C-Q, a shortfall in the HMM method is easily overcome. The inclusion of auto-correlation in the HMM requires that the independent variable have no gaps; and this is more likely in precipitation records than in streamflow records. Whilst comprehensive precipitation data over decades and methods to extract such data for specific catchments is available, the same cannot be said for the availability of streamflow data.

2.1.2 *Site Selection*

In Victoria , 42 DELWP sites had high frequency EC data collected at a minimum of 2-hours apart between readings. The Millennium Drought (~1995-2009) provides a natural experiment over which the effects of such extreme hydrological events could be studied. Thus, this period (~1990-2020) was the focus of this study. The following three criteria were used to select sites for investigation. Firstly, heavily regulated catchments were excluded from the study. Sites with dams/offtakes that accounted for less than 5% of mean annual flows were included still. All selected sites were then required to have at least 1,000 readings a year, for years in which Q and EC data were collected, to confirm that the data collected was of the 'frequent' type. All sites passing these criteria were selected, should Q and EC data be fully available each year over the period of 1991-2015, representing approximately 5 years pre- and post-drought. For all other sites failing the last point above, these were still selected if there were less than 2 years of data missing between 1991-2015, provided there was 2 years or more of data either before 1991 or after 2015.

Out of the 42 sites, only 8 sites successfully met the criteria above. For these 8 sites, point EC data was aggregated to monthly TDS loads. Where there were more than 3 days per month in which there was no EC or streamflow reading, that month's TDS flux load was set to N/A. As HMMs can handle gaps in the predicted variable, this approach was preferable to introducing significant errors of 10% or more in TDS load estimation. The proportion of invalid (thus missing) monthly TDS data was up to 50% at some sites.

2.2 *Hidden Markov Modelling*

2.2.1 *Model Set Up*

Salinity (monthly or seasonal) data at time t was modelled as originating from one of two distinct steady states i. Generalising Peterson et al. (in-review), the following modification is proposed in Equation 1:

$$\widehat{{}_tC_i} = \alpha_i + \beta P_t + \phi_1 \cdot \widehat{{}_{t-1}C_i} + \phi_2 \cdot \widehat{{}_{t-2}C_i} + \phi_3 \cdot \widehat{{}_{t-3}C_i} \quad (1)$$

where the conditional mean for the distribution $\widehat{{}_tC_i}$ is modelled as a function of the predictor variable P_t for rainfall, and lag-1 up to lag-3 autocorrelation terms. The parameter α_i is state-dependent to allow for shifting salinity relationships, whilst β is state-independent and defines the transportation ability of rainfall on salt loads in the case of flux data. Φ_1, Φ_2, Φ_3 are state-independent parameters for lag-1, lag-2, and lag-3 autocorrelation, respectively. The variance of the predicted salt load's distribution varies with state but does not vary with time.

Following Peterson et al. (in-review) Equations 4-6, 9 and 12, the state $C^{(t)}$ at time t is a Markov function of the two-state transition matrix Γ_2, for which these probabilities are time-invariant, shown in Equation 2 below:

$$\Gamma_2 = \begin{vmatrix} p_{11} & 1-p_{11} \\ 1-p & p_{22} \end{vmatrix} \quad (2)$$

where the probability of the state at t, transitioning from $C^{(t-1)}$ to $C^{(t21)}$ is determined by Equation 3:

$$\mathrm{p}_{12} = \mathrm{pr}\left(\mathrm{C}_2^{(t)} | \mathrm{C}_1^{(t-1)}\right) \quad (3)$$

In total, 32 different models were calibrated (16 single models and 16 two-state models), where the means, auto-regression terms, and/or variances were allowed in varying combinations to differ in each state and/or month. Then, as in Peterson et al. (in-review) Equation 17, the likelihood function, shown here as Equation 4:

$$L_T = \delta P(x_1)\Gamma P(x_2) \dots \Gamma P(x_T)1' \tag{4}$$

for an initial state distribution of δ, emission probabilities P obtained from a Gamma distribution of model probability densities, for all *t* timesteps, is optimised.

This calibration was done using a differential evolution genetic algorithm to identify the global optima through maximum likelihood estimation, following Peterson et al. (in review). During each optimisation, 25 parameter sets times the number of parameters were randomly sampled and allowed to converge to a maximum on the optimisation surface for up to 10,000 generations. For each of the 32 models, the optimisation was done at least ten times, to ensure the global optima was identified. All modelling was done by extending the hydroState R package (Peterson et al., in review).

2.2.2 *Model Selection*

For each of the calibrated 32 models, the Akaike Information Criterion (Akaike, 1976) was calculated. The model with the lowest AIC, where $AIC = 2N + 2L_{TS}$ for *N* number of model parameters, was chosen. Should a two-state model be chosen as the optimal model regardless of the adverse effects of its many more parameters on the calculated AIC, in contrast to one-state models, this suggests that in-stream salinity behaviour may be better explained through the existence of more than one steady state.

2.2.3 *Model Validity*

The validity of the models was tested by checking that the pseudo-residuals were normally distributed (Zucchini and MacDonald, 2009). This was assessed using pseudo-residual plots and the Shapiro-Wilk test (alpha = 0.05). For models to be accurate, the pseudo-residuals must be normally distributed. Viewing the autocorrelation plots of the pseudo-residuals (Figs. 3A and 3B) also enables determination if the model has performed sufficiently well. If minimal autocorrelation is seen to carry into future time-steps, it signifies that errors do not accumulate into the future and that no information is 'leftover' and unmodelled.

Figure 3A gives an example of pseudo-residual analysis showing poor model performance. The auto-correlation of the pseudo-residuals are outside of acceptable bounds for many consecutive timesteps, indicating that significant information from the data has not been included in the model. In contrast, Figures 3B-D give an example of relatively strong pseudo-residuals. Particularly, Figure 3B shows minimal serial correlation of the pseudo-residuals, which drop to below 0.2 by the second lag, signifying that inaccuracies in the model at one time

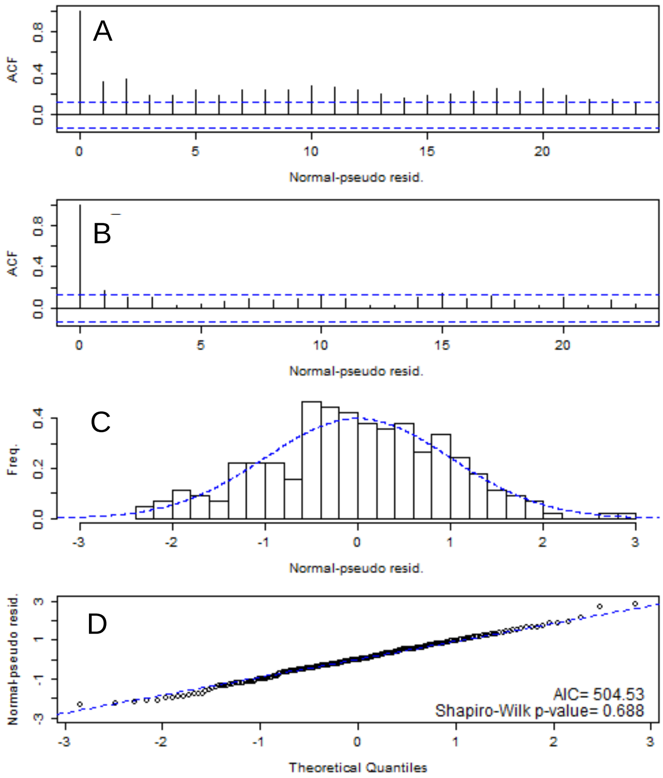

Figure 3. Pseudo-residual analysis of HMM model, (A) at Site 405212, (B)-(D) at Site 406213. (A) aand (B) Auto-correlation functions of the normal pseudo-residuals. (C) Histogram of the normal distribution. (D) Quantile-quantile plots of normal pseudo-residuals, AIC and Shapiro-Wilk p-values noted.

step have minimal effect on subsequent timesteps. Figure 3C and D also show normally distributed pseudo-residuals and a Shapiro-Wilk p-value of 0.688, respectively. This p-value, being larger than 0.05, confirms that the pseudo-residuals are normal, and that model performance is strong.

2.2.4 *Model Results*

The Viterbi algorithm (Viterbi, 1967) was then used to identify from the Markov Chain of probabilities, what the most probable sequence of states observed was. From Viterbi plots of the sequence of Markov states, the classified time series of salinity data could also be obtained, where periods in which salinity data belonged to different states were distinguished. This classified time series could then be visually compared to other time-series data, such as the same output for rainfall-runoff time series, for investigation of timing and general patterns.

2.2.5 *Sensitivity to Input Data*

The sensitivity of the modelling process to different proportions of missing data was first investigated to check the importance of quality of data coverage. We were uncertain whether gaps in the data would have a large effect on the model results obtained. A sensitivity analysis was carried out, using the data of the site with the least number of gaps (Site 234201).

Table 1. Distribution of skipped readings in simulated data.

Case	*Number of days missing per month*	*Number of months in study period*	*Data coverage over studied period (%)*
1	*≤ 3*	*245*	*75%*
2	*≤ 6*	*270*	*83%*
3	*≤ 9*	*286*	*88%*
Original	*≤ 31*	*324*	*100%*

The original raw data set for this site covered a period of 324 months, where only 305 of those months were considered to have full TDS monthly loads, due to having less than 3 days of skipped readings in these months (94% data coverage). Raw EC data values were then deleted, simulating the distribution of skipped readings in an average 'poor' quality data set. This resulted in three data input cases for analysis (Table 1).

3 RESULTS

3.1 *Sensitivity to Input Data*

HMM models were calibrated for aggregated monthly TDS loads, allowing up to 3 days of skipped readings per month before the whole month was nullified (75% data coverage – Case 1, as per Table 1). Additional sensitivity analysis modelling was undertaken, now allowing up to 6 and 9 days missing per month instead before it was nullified (Case 2 and 3, respectively). In all cases, two Viterbi salinity states were identified by the best model (Fig. 4), as in results using the original data (Fig. 5).

For Case 1, with the highest percentage of missing data, only a few months at the start of the study period were in the 'normal' state , before a long sustained shift to a low salinity state occurred, similar to that observed in the original data (Fig. 5). Results obtained for Cases 2 and 3 were similar to each other, where the timings of the salinity state changes were different to that observed in the original data. These results suggest that even with a higher number of missing months, HMMs are still sufficiently robust and able to identify the general trend in terms of timing of changes in states. However, the results obtained from Case 2 and 3 indicate high sensitivity to accurate aggregation of values for each monthly time step. Thus, months with lower data frequency could not be included in the modelling process without greatly compromising the results.

3.2 *Salinity States*

Monthly TDS loads at eight sites across Victoria were modelled using precipitation data. At all 8 sites, the best model chosen (by AIC) was a two-state model (Fig. 6). Compared to 1993 as a base year, in which in-stream salinity was in a 'normal' state, the second alternative stage identified was a 'low' salinity state at all sites. At 6 out 8 study sites, salinity shifted to a low state only after the onset of the Millennium Drought (Fig. 6).

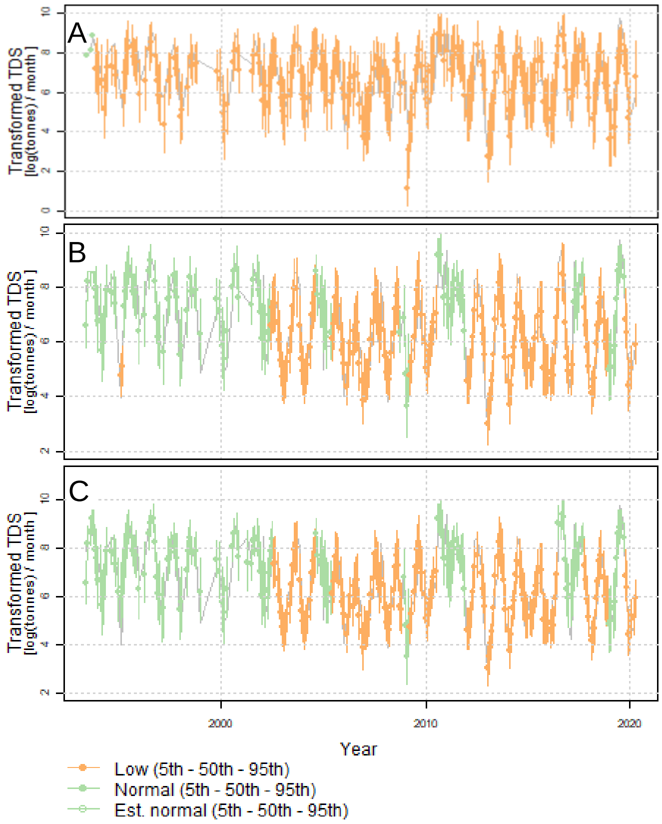

Figure 4. Time series results for sensitivity to input data analysis showing of observed total dissolved solid flux loads. HMM states (vertical bar extends from the 5th to 95th percentiles, point denotes median) from Viterbi algorithm decoding and the estimated transformed TDS load for the catchment in a low state. (A) Case 1, 75% coverage, months with >3 day gaps excluded (B) Case 2, 83% coverage, months with >6 day gaps excluded. (C) Case 3, 88% data coverage, months with >9 days gaps excluded.

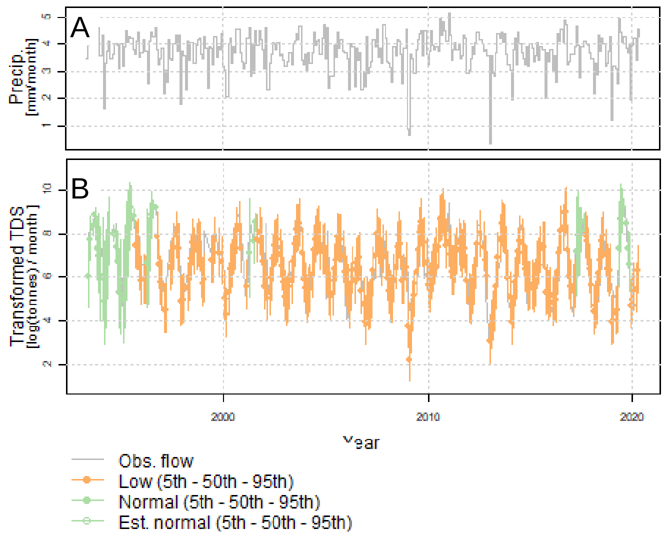

Figure 5. Time series results for Site 234201. (A) Monthly precipitation. (B) Time series of observed total dissolved solid flux loads. HMM states (vertical bar extends from the 5th to 95th percentiles, point denotes median) from Viterbi algorithm decoding and the estimated transformed TDS load where the catchment was in a low state.

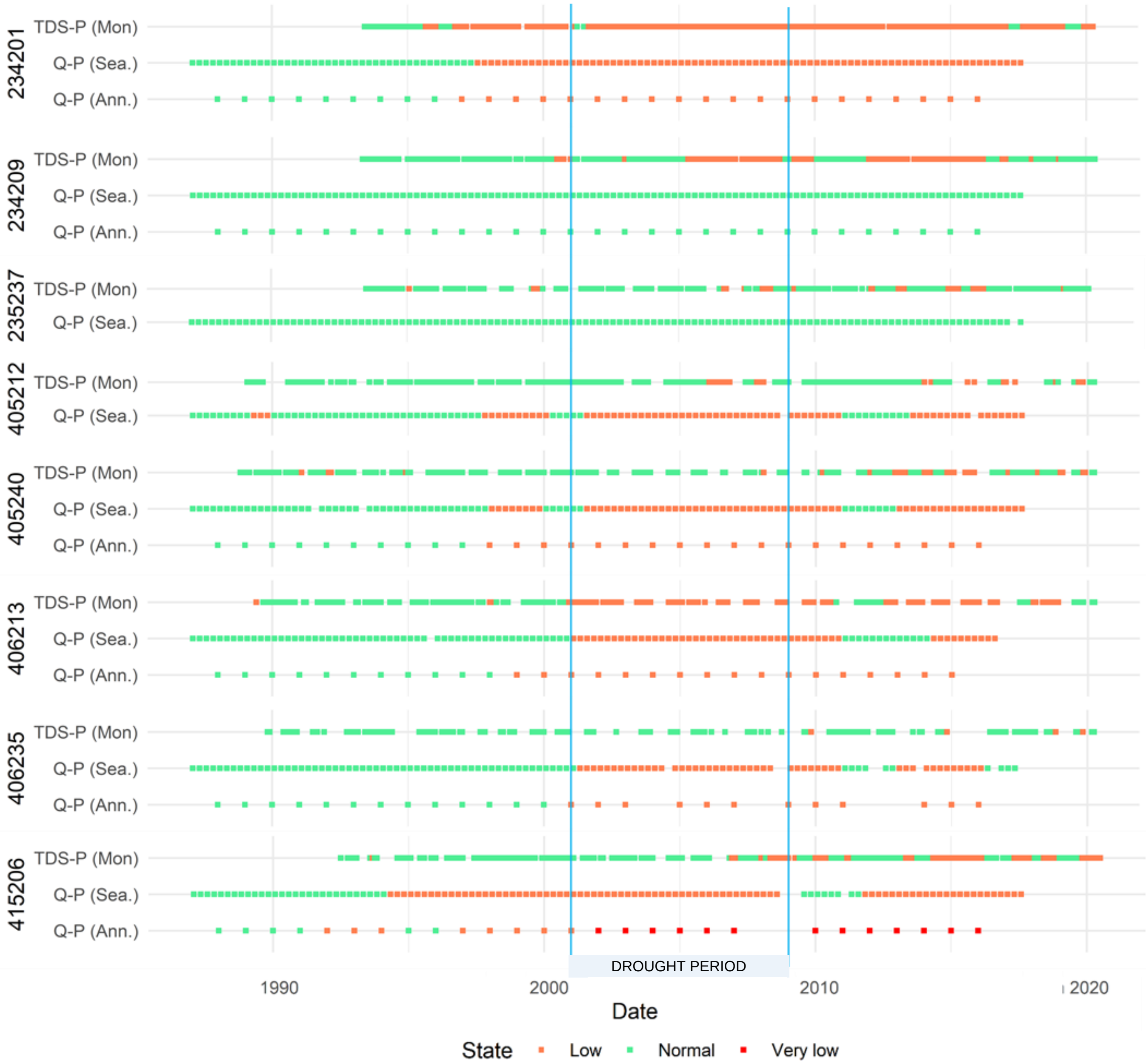

Figure 6. Summary of Viterbi states identified by HMM for each site through use of: (i) total monthly dissolved solid loads i.e. salinity, predicted by total monthly precipitation; (ii) total annual streamflow predicted by total annual precipitation; (iii) total seasonal streamflow predicted by total seasonal precipitation. Precipitation-streamflow Viterbi states courtesy of Peterson et. al. (in review). The period of the Millenium Drought, from approximately 2001-2009, is delineated in blue.

Low salinity states mean that historical TDS loads during such periods were less than that expected, even if precipitation remained constant. This indicates a change in the relationship between salinity and precipitation. Figure 6 summarises the Viterbi states at each of the eight sites obtained though HMM analysis of monthly TDS predicted by monthly total precipitation (shown in the first row of each dot point grouping per site). For five of the six sites in which salinity state changes occurred after the start of the Millennium Drought, it appeared that whilst two-state models better reflected in-stream salinity responses, the switch to the alternate 'low' state occurred sporadically and intermittently. Upon visual inspection, only 3 sites (234201, 234209, and 406213) showed relatively long-term periods over which in-stream salinity switched to a 'low' state.

Site 234201 showed the longest sustained shift into a low salinity state, throughout almost all of the study period. A full example of the Viterbi states identified for this site was shown in Figure 5. Similarly, Site 406213 switched into a low salinity state at the start of the drought, briefly changing back to a normal state in 2001, before re-entering a low salinity state, which was sustained until 2010. Here a brief switch back into the 'normal' salinity state was again observed. This switchback appeared to coincide with the La Niña event which produced above average precipitation. After this period, the catchment continued to shift in and out of the low salinity state.

3.3 *Comparison to rainfall-run-off states*

Aside from the summary of Viterbi states observed in monthly salinity data, Figure 6 also summarises the results of HMM analysis carried out at the seasonal and annual scale, for total streamflow predicted by total precipitation for that period (Peterson et al., in review). Half of the sites showed a shift to a low salinity state only occurring after the catchment had shifted into a low runoff state for some time. Two sites (234209, 235237) showed changes to low salinity states but not runoff states, where changes to a low salinity state occurred much later than the initial onset of the Millennium Drought. Two sites (234201, 406213) showed changes to a low salinity state beginning at about the same time a shift to a low runoff state occurred, at the start of the Millennium Drought.

3.4 *Model Validity*

Having determined that two state models better explain stream salinity than single-state models, the performance of each sites' best HMM model was also examined, with key features summarised in Table 2. Only 3 models had Shapiro-Wilk values above 0.05, indicating that model performance was generally poor. This is also observed in the ACFs, where serial correlation of errors remains high throughout many lag intervals. The percentages of missing monthly TDS loads over the full study period at each site is also shown in Table 2, where such data was nullified due to having skipped more than 3 days of readings in that month (signifying incomplete TDS loads, as per Section 2.1.2 requirements). Minimal correlation is seen between higher proportions of missing TDS data and reduced model performance.

Table 2. Summary of pseudo-residual analysis.

Site	***Months Missing (%)***	***Shapiro-Wilk p-value***	***Lag where ACF <0.4***	***Lag where ACF <0.2***
234201	*6%*	*0.548*	*1*	*7*
234209	*10%*	*0*	*1*	*1*
235237	*24%*	*0.013*	*2*	*2*
405212	*32%*	*0.022*	*1*	*22*
405240	*39%*	*0.434*	*1*	*1*
406213	*40%*	*0.688*	*1*	*2*
406235	*57%*	*0.021*	*5*	*10*
415206	*24%*	*0.03*	*2*	*12*

4 DISCUSSION

The analysis conducted in this study indicates that multiple steady states of stream salinity regimes exist. If years prior to drought are considered 'normal' , the years following a drought are likely to see shifts into 'low 'salinity states, reflecting a lower total volume of salts entering our rivers per unit rainfall. This salt comes from a variety of sources, which are mobilised, then transported into our rivers in different ways, by both surface and subsurface flows (Lintern et al., 2018a). Due to the non-urban nature of the catchments studied, deposits of dissolved salt in the soil and bedrock are the likely major sources of in-stream salinity (Lintern et al., 2018a). These catchments also have high levels of dryland agricultural activity (averaging 80%), which could lead to considerable mobilisation of dissolved salts from the surface, from rising saline groundwater tables depositing salts closer to the ground surface, which are then washed out during rain events.

Typically during drought, in-stream salinity in terms of concentrations (as electrical conductivity) has been found to increase, due to lower mobilisation and thus dilution ability, as a result of reduced precipitation, which leads to reduced flushing in catchments (Mosley, 2015). Since the HMMs were modelled using salt loads, rather than concentrations, it is possible that 'low' Viterbi states (i.e., decrease in salt load) are be driven by a decrease in streamflow, rather than a decrease in salt concentration. It is likely then, given that most salt comes from groundwater, lower salt volumes could be due to a reduction in highly saline groundwater baseflow contribution. At the studied sites, the mean baseflow index for these catchments were typically 0.5 and above (Lintern et al., 2018b), indicating that streamflow in these catchments would indeed be influenced by baseflow processes. In catchments well connected to groundwater, in-stream salt concentrations may instead not vary much during drought (Wilbers et al., 2009). Further HMM analysis could be repeated on EC-Q, to clarify whether the shift to the low salinity state was driven by shifts in either EC concentration or streamflow.

Analysing the timing of the salinity state changes compared to runoff state changes provides further clues. Overall, shifts to a low salinity state only occurred some years after a shift to a low runoff state had already occurred. Groundwater acts as a catchment's "memory" (Hughes et al., 2012), where for groundwater storage processes, and thus salt volume processes, the effect of external changes is buffered by the conditions of previous years. This may explain the lag observed in the salinity states behind the runoff states, indicating that typically, changes in baseflow could be the reason why salinity state changes were also observed. If surface flow mechanisms were the more likely cause of the changes in salinity states, we would instead expect more sites to show changes to a lower salinity state sooner. This was observed at only two sites (234201, 406213), where in-stream salinity may be more affected by surface flows than baseflows.

The lag between runoff and salinity state changes strongly indicate that changes in baseflow processes do not cause the decreased runoff observed in catchments during and after drought. Instead, it suggests that changes in baseflow occur because of reduced runoff occurring beforehand. Perhaps, at first occurrence of drought, baseflow continues to contribute to streamflow as before, whilst other catchment processes such as evapotranspiration adjust their response, leading to decreased runoff. During this time, relatively more groundwater is discharged than recharged, up until a point in which groundwater discharge reacts by reducing too, as per Peterson et. al (in review). Furthermore, the shift to a low salinity state continues to occur even past the Millennium Drought, reflecting the non-recovery of runoff also observed by Petersen et al. (in review). Further time series records and analysis are required to investigate if this is a consequence of the 'memory' function of groundwater, meaning that the time had not come yet for baseflow processes to normalise at the end or the drought, or if instead, such results further confirm the limited resilience of our catchments and their processes.

Finally, the pseudo-residuals analysis shows that model performance could be further improved, as information remains in the errors not included in the model structure. Inclusion of this information is unlikely to change the conclusions of the existence of multiple steady states but may improve the fit with observed data. Box-Cox transforms may also be implemented to manage heteroscedasticity. Future HMMs may make use of more complex model structures, such as the through the development of multivariate HMMs. This may be done similarly to Spezia et al. (2011), where instead of multivariate analysis comprising of only water quality variables, joint analysis of both water quality and non-water quality data, such as joint analysis of salinity and streamflow data as their own variables, may be explored.

5 CONCLUSION

Riverine salinity fluxes are better modelled by HMMs through the existence of multiple steady states, than single states. Six out of eight catchments showed shifts to low salinity states a few years after the start of the Millennium Drought. At four of these sites, these changes came after a shift to a low runoff state had already occurred. Catchments may switch to low salinity states intermittently but have continued to do so since the drought occurred. A decrease in groundwater discharge is likely to have caused less total salt loads in these periods. These different states provide insight into the changes in catchment hydrological processes during the Millennium Drought, specifically, that the reduction in groundwater discharge is not a causal factor leading to missing and unrecovered runoff in these catchments. Further investigation through the application of, and the development of, multivariate HMM analysis may be able to improve model performance, therefore improving predictions of the exact durations of these state changes.

The implications of these findings go beyond in-stream salinity, likely affecting other water quality constituents, which follow similar processes in our catchments. This highlights the limited resiliency of our catchments in terms of water quality processes and adds to growing knowledge that catchments can have more than one steady state, in general. With climate change increasing the severity and frequency of extreme events able to cause state changes, further understanding of the causes of water quality state changes is required to ensure proper catchment management at the most potentially vulnerable sites.

6 REFERENCES

Akaike, H. (1976), "Canonical Correlation Analysis of Time Series and the Use of an Information Criterion", Mathematics in Science and Engineering, pp. 27–96.

Bieroza, M.Z., Heathwaite, A.L., Bechmann, M., Kyllmar, K. and Jordan, P. (2018), "The concentration-discharge slope as a tool for water quality management", Science of the Total Environment, Vol. 630, pp. 738–749.

Cañedo-Argüelles, M., Kefford, B.J., Piscart, C., Prat, N., Schäfer, R.B. and Schulz, C.J. (2013), "Salinisation of rivers: An urgent ecological issue", Environmental Pollution, Vol. 173, pp. 157–167.

Fisher, J.R., Dvorak, B.I. and Admiraal, D.M. (2016), "Pollutant Load Estimates Using Regression Models with In-Stream Measurements", Journal of Environmental Engineering, Vol. 142 No. 3, p. 04015081.

Godsey, S.E., Hartmann, J. and Kirchner, J.W. (2019), "Catchment chemostasis revisited: Water quality responds differently to variations in weather and climate", Hydrological Processes, John Wiley and Sons Ltd, Vol. 33 No. 24, pp. 3056–3069.

Hirsch, R.M., Moyer, D.L. and Archfield, S.A. (2010), "Weighted regressions on time, discharge, and season (WRTDS), with an application to Chesapeake Bay River inputs", Journal of the American Water Resources Association, John Wiley & Sons, Ltd, Vol. 46 No. 5, pp. 857–880.

Holling, C.S. (2013), "Resilience and stability of ecological systems", The Future of Nature: Documents of Global Change, Yale University Press, Vol. 4 No. 1, pp. 245–256.

Hughes, J.D., Petrone, K.C. and Silberstein, R.P. (2012), "Drought, groundwater storage and stream flow decline in southwestern Australia", Geophysical Research Letters, Vol. 39, No. 3.

Kho, N., Lintern, A., Guo, D., Liu, S. and Duvert, C. (2020) "Using historical data to identify future water quality trends at a regional scale", EGU General Assembly 2020, Online, 4–8 May 2020, EGU2020-4308, available at: https://doi.org/10.5194/egusphere-egu2020-4308.

Knapp, J.L., Freyberg, J. von, Studer, B., Kiewiet, L. and Kirchner, J. (2020), "Concentration-discharge relationships vary among hydrological events, reflecting differences in event characteristics", Hydrology and Earth System Sciences Discussions, pp. 1–27.

Lintern, A., Webb, J.A., Ryu, D., Liu, S., Bende-Michl, U., Waters, D., Leahy, P., et al. (2018a), "Key factors influencing differences in stream water quality across space", Wiley Interdisciplinary Reviews: Water, Wiley, Vol. 5 No. 1, p. e1260.

Lintern, A., Webb, J.A., Ryu, D., Liu, S., Waters, D., Leahy, P., Bende-Michl, U. and Western, A.W. (2018b). "What are the key catchment characteristics affecting spatial differences in riverine water quality?", Water Resources Research, Vol. 54 No. 10, pp. 7252–7272, available at: https://doi.org/10.1029/2017WR022172.

Minaudo, C., Dupas, R., Gascuel-Odoux, C., Roubeix, V., Danis, P.A. and Moatar, F. (2019), "Seasonal and event-based concentration-discharge relationships to identify catchment controls on nutrient export regimes", Advances in Water Resources, Vol. 131, available at: https://doi.org/10.1016/j.advwatres.2019.103379.

Moatar, F., Abbott, B.W., Minaudo, C., Curie, F. and Pinay, G. (2017), "Elemental properties, hydrology, and biology interact to shape concentration-discharge curves for carbon, nutrients, sediment, and major ions", Water Resources Research, Blackwell Publishing Ltd, Vol. 53 No. 2, pp. 1270–1287.

Mosley, L.M. (2015), "Drought impacts on the water quality of freshwater systems; review and integration", Earth-Science Reviews, vol. 140, pp. 203–214.

Musolff, A., Schmidt, C., Selle, B. and Fleckenstein, J.H. (2015), "Catchment controls on solute export", Advances in Water Resources, Vol. 86, pp. 133–146.

Peterson, T.J., Argent, R.M., Western, A.W. and Chiew, F.H.S. (2009), "Multiple stable states in hydrological models: An ecohydrological investigation", Water Resources Research, Vol. 45 No. 3.

Peterson, T.J., Peel, M.C. and Saft, M. (2019), "Catchment Drought Recovery : A statistical analysis of hydrological resilience throughout Victoria", Abstracts: 23rd International Congress on Modelling and Simulation, The Modelling and Simulation Society of Australia and New Zealand Inc, Canberra, p. 483.

Peterson, T.J., Saft, M., Peel, M.C. and John, A. (in review). "Watersheds may not recover from droughts", pp. 1–14.

Peterson, T.J. and Western, A.W. (2014), "Multiple hydrological attractors under stochastic daily forcing: 1. Can multiple attractors exist-", Water Resources Research, Vol. 50 No. 4, pp. 2993–3009.

Peterson, T.J., Western, A.W. and Argent, R.M. (2012), "Analytical methods for ecosystem resilience: A hydrological investigation", Water Resources Research, Vol. 48 No. 10, available at: https://doi.org/10.1029/2012WR012150.

Rousseeuw, K., Poisson Caillault, E., Lefebvre, A. and Hamad, D. (2015), "Hybrid hidden markov model for marine environment monitoring", IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, Institute of Electrical and Electronics Engineers, Vol. 8 No. 1, pp. 204–213.

Rue, G.P., Rock, N.D., Gabor, R.S., Pitlick, J., Tfaily, M. and McKnight, D.M. (2017), "Concentration-discharge relationships during an extreme event: Contrasting behavior of solutes and changes to chemical quality of dissolved organic material in the Boulder Creek Watershed during the September 2013 flood", Water Resources Research, Blackwell Publishing Ltd, Vol. 53 No. 7, pp. 5276–5297.

Saft, M., Western, A.W., Zhang, L., Peel, M.C. and Potter, N.J. (2015), "The influence of multiyear drought on the annual rainfall-runoff relationship: An Australian perspective", Water Resources Research, Blackwell Publishing Ltd, Vol. 51 No. 4, pp. 2444–2463.

Scheffer, M. and Carpenter, S.R. (2003), "Catastrophic regime shifts in ecosystems: Linking theory to observation", Trends in Ecology and Evolution.

Shafiei, F. and McLoughlin, P.D. (2017), "Change in nutrient loading pattern due to coupled effect of change in concentration and hydroclimatic forces", Journal of Freshwater Ecology, Taylor and Francis Inc., Vol. 32 No. 1, pp. 773–792.

Spezia, L., Futter, M.N. and Brewer, M.J. (2011), "Periodic multivariate normal hidden markov models for the analysis of water quality time series", Environmetrics, John Wiley & Sons, Ltd, Vol. 22 No. 3, pp. 304–317.

Suchetana, B., Rajagopalan, B. and Silverstein, J.A. (2019), "Investigating regime shifts and the factors controlling Total Inorganic Nitrogen concentrations in treated wastewater using non-homogeneous Hidden Markov and multinomial logistic regression models", Science of the Total Environment, Elsevier B.V., Vol. 646, pp. 625–633.

Thyer, M. and Kuczera, G. (2003), "A hidden Markov model for modelling long-term persistence in multi-site rainfall time series. 2. Real data analysis", Journal of Hydrology, Elsevier, Vol. 275 No. 1–2, pp. 27–48.

van Dijk, A.I.J.M., Beck, H.E., Crosbie, R.S., de Jeu, R.A.M., Liu, Y.Y., Podger, G.M., Timbal, B. and Viney, N.R. (2013), "The Millennium Drought in southeast Australia (2001-2009): Natural and human causes and implications for water resources, ecosystems, economy, and society", Water Resources Research, Vol. 49 No. 2, pp. 1040–1057.

Viterbi, A.J. (1967), "Error Bounds for Convolutional Codes and an Asymptotically Optimum Decoding Algorithm", IEEE Transactions on Information Theory, Vol. 13 No. 2, pp. 260–269.

Walker, B., Holling, C.S., Carpenter, S.R. and Kinzig, A. (2004), "Resilience, adaptability and transformability in social-ecological systems", Ecology and Society, Vol. 9 No. 2, available at: https://doi.org/10.5751/ES-00650-090205.

Wilbers, G.J., Zwolsman, G., Klaver, G. and Hendriks, A.J. (2009), "Effects of a drought period on physico-chemical surface water quality in a regional catchment area", Journal of Environmental Monitoring, Vol. 11, Vo. 6, pp. 1298-1302.

Zucchini, W. and MacDonald, I.L. (2009), Hidden Markov Models for Time Series, Hidden Markov Models for Time Series.

MANAGEMENT STATEMENT

My FYP experience was full of ups and downs, but now at the end of it, I feel positively about the unit, and proud of my work. At times, I second guessed myself for taking on a statistics project, especially when I struggled to understand concepts even after putting in lots of effort, particularly in the beginning. Now, I am glad to have challenged myself and learnt new skills. I have gained so much respect for researchers – it is such hard work, both mentally and emotionally.

Reflecting on my experience, if I could go back I time, I would have calmed myself down and not gotten caught up in trying to make things ‘perfect’. I was initially so overwhelmed. For example, the dataset I wanted to use didn’t work out for the analysis method, and then I had to hunt down the data type I needed, and there was much less of it, and it had to be further prepared as well. I’ve learnt that real world data is not perfect, and even though it is important to clean it up properly and get it into the state you need, this can take so much time – to figure out how you are going to write to code to do it, actually writing the code and debugging it, and then checking that the results are right. It was important not to underestimate this and make time for it appropriately, but at some point I learnt to cut my losses or else I would not get to the other, more difficult parts of analysis. The same went for doing literature reviews – there is so much knowledge out there, some of it so hard to understand, and yet still so much more to be discovered. I learnt to be patient with myself and read a little day by day to not get overwhelmed, and to not expect to find all the answers all the time.

Furthermore, whilst I did keep a physical logbook of progress and meeting notes, my writing became messier the later into the semester it got, until I could not even read it. Also, all my data and results were on my computer and could not be simply written down. Therefore, my results and notes were all over the place and disorganised, which was a real detriment to me when I was trying to look back on the things I had noted or done. Furthermore, in terms of general organization, I had 20 different versions of the same code, updated with names as ‘V2’ and ‘V3’ and so on in a folder for a particular task. However, I would also have multiple versions of that folder, and sometimes I would revert to using older versions if newer versions ended up being wrong. Therefore, it was hard to keep track of exactly which file was the right one. In my future projects, I would implement much better organization. I would keep all my notes in Microsoft OneNote instead (which I only ended up doing towards the very end of the project). This would have been much more helpful as I could paste all sorts of media, then organise notes into different sections. I would also spend more time managing my documents and keeping notes on versions, instead of dumping files all over the place, and properly archive retired files.

In terms of what went well, time management-wise, I began working early in the semester, which was a good idea as even the smallest efforts built up over time. I probably could have gotten more done during the break between Semester 1 and 2, but I needed time to recharge (thought I probably rested *too* much). Also, whilst learning to code in R was something relatively new, after a lot of StackOverflow, googling how to use different packages, and lots of practice, I have since gained new skills coding in R. I also found resilience in myself – so many times, so many things just did not work, like having to re-prep data over and over because there was always something wrong with the method. Even with a logbook full of angry red cross marks, I am glad to have not given up. Finally, I learnt the importance of good mentors and support systems. I was extremely lucky to be working under Tim, my official supervisor who provided so much motivation and guidance; and Anna, who previously supervised me during my summer research project, who still came back after maternity leave to give good advice and pointers. I am so grateful for their belief in me, providing direction whenever I was lost and stressed (which was frequently), and for even taking my frantic weekend phone calls. What a semester it has been, glad to have made it to the end!