HIERARCHICAL BAYESIAN SPATIO-TEMPORAL MODELLING FOR CONCEPTUAL PROCESS UNDERSTANDING

Joshua A. Simmons, DARE, The University of Sydney, joshua.simmons@sydney.edu.au
Kristen D. Splinter, UNSW Sydney, k.splinter@unsw.edu.au
R. Willem Vervoort, DARE, The University of Sydney, willem.vervoort@sydney.edu.au

INTRODUCTION

Shorelines can be seen as representations of the constant interaction between hydrodynamic forces from the ocean and available sediment supply. The modelling of shoreline movement over short to medium timescales remains an active area of research owing to the difficulty of accounting for all the different factors influencing these dynamics. These include both cross-shore and alongshore processes acting at a range of timescales from individual waves to multi-decadal trends in waves and water levels. While many models to date have focused on either alongshore or cross-shore processes, the connectivity between adjacent locations requires a more inclusive approach. Along complex coastlines, coastal management requires flexible methodologies for pulling apart the signal considering a range of sources at any given location.

There has previously been a paucity of consistently sampled multi-decadal datasets, barring a few well monitored sites (e.g., Narrabeen Beach). The growing availability of freely available satellite imagery has led to a surge of shoreline data (Vos et al., 2019). The timespan (multi-decadal) and spatial coverage (globally) of this data has meant a step change in available information for most locations. This has opened the door to a number of studies leveraging these data to inform quantitative coastal management (Kennedy et al., 2023). As these data become more widely used, it is important to have approaches that extend beyond regression trends and attempt to quantify the complex underlying signals present (Luijendijk et al., 2018). This is particularly important when considering climate change impacts into the future, as the assumption of past and future conditions being similar cannot be taken for granted. More nuanced models are needed that can consider which shoreline response signals are likely to persist and which may change.

To address these types of questions, data-driven modelling approaches have the potential to be particularly beneficial given the scope of satellite datasets. With the wealth of data available it can be tempting to employ more black-box machine learning models such as neural nets and other deep learning approaches. These modelling approaches have shown considerable predictive skill when hindcasting storm events and shoreline movement (Simmons et al., 2019), however much work still must be done to make them interpretable for coastal management. Approaches such as Bayesian Networks provide much more transparent data-driven tools, however these simplify over the complex and often nonlinear interactions that quantify the relationships described in the models.

It is important to be able to explicitly embed existing causal knowledge within a flexible and interpretable framework. When learning the regression coefficients of the model implied by our causal understanding, it is also important to quantify uncertainty. Shoreline signals are complex and care must be taken to understand the degree to which different parameter estimates are supported by the data, especially when considering implications for future management decisions. Bayesian inference applied to regression modelling has the potential to fulfill these requirements. Approaches such as Bayesian spatio-temporal models have been around for some time (Wikle et al., 1998). However, the growing availability of fast and efficient samplers for inference, and the maturity of probabilistic programming libraries have made these models increasingly feasible.

METHODOLOGY

This study uses data from Geoscience Australia's DEA coastlines dataset (Bishop-Taylor et al., 2021) which is comprised of yearly shoreline positions derived from optical satellite imagery at 30m intervals around the Australian coastline from the late 1988 to the present. Shorelines are identified using a sub-pixel method to delineate the waterline from a composite of cloud free and tidally masked Landsat images (30m resolution) across each year. These yearly shoreline positions have been extracted for this study at a number of sites exhibiting different coastal processes, including Narrabeen Beach in Sydney and Trial Bay Beach near South West Rocks in NSW, Australia. To ensure the number of spatial data points were manageable at each site, yearly shoreline positions were averaged alongshore considering of the scale of coastal processes being examined. The final data consists of average shorelines every 150m and 300m at each site, respectively.

Taking a Bayesian approach, the data can be described by a process model (μ) plus a normally distributed error term with a mean of 0 and a variance of σ^2.

$$y_{i,t} = \mu_{i,t} + \epsilon \quad \epsilon \sim \mathrm{N}(0, \sigma^2)$$

A simple modelling approach is then to assume that change in shoreline at each time step can be estimated using a linear regression on the variables of interest, plus random processes to account for spatio-temporal dependencies. For example at South West Rocks (leaving aside random processes) we can model the shoreline at each timestep (t) and alongshore location (i) to be:

$$\mu_{i,t} = y_{i,t-1} + \beta_{lagged,i} X_t + \beta_{E,i}\, E_t + \beta_{trend,i}$$

where the change in shoreline is composed of a trend term (β_{trend}), a term (β_E) representing the response to yearly cumulative offshore storm energy (E_t), and a term (β_{lagged}) measuring the contribution of a lagged (up to 3 years) signal from Laggers point (X_t). The Laggers point in the signal is taken as the closest profile to the headland at the east of the site (see Figure 1), given the predominant flow of sediment from south to north along the NSW coast. The lagged signal term (β_{lagged}) therefore represents the degree to which pulses of longshore sand transport around the headland influence the shoreline position at the various locations along the beach.

Given that coastal processes usually act over spatial scales beyond a single alongshore location, a hierarchical modelling approach can be used to promote information sharing between parameters. Ideally the Bayesian prior on the parameters of the model represents our understanding that the contribution of each of the forcing terms should be similar the closer the two profiles are in the alongshore. To this end hierarchical gaussian process priors are used, which serve to regularise the parameters learned from the data. The Bayesian approach quantifies uncertainty in the parameter values, which helps to diagnose where multiple factors could be contributing to shoreline movement and it is unclear which is dominant.

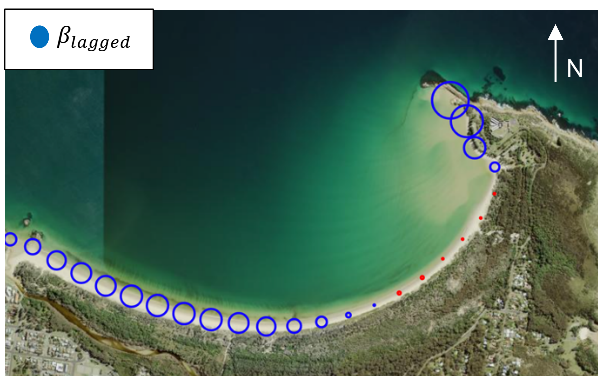

Figure 1 - Coefficient (β_{lagged}) on the lagged Laggers Point signal showing the variable degree of connectivity with headland bypassing processes. The size of the blue circle indicates the positive magnitude (and red the negative) of the coefficient and as such the contribution of the lagged bypassing to shoreline movement at that location. *Imagery: NSW Six Maps.*

RESULTS

Figure 1 shows the β_{lagged} term from the model (maximum over all lags) which reveals the high degree of connectivity on the eastern end of the beach. Here sediment pulses moving around the headland move southward over relatively short timescales. Towards the centre of the beach, shoreline movement becomes less correlated to these headland bypassing events. Generally, the model is less skilful in predicting the yearly shoreline series here, an indication of potential missing drivers in the model that determine longshore transport in this location. Evidence of subaqueous sediment transport across the bay is apparent as the correlation to the Laggers point signal increases to the west of the bay, though with a longer lag from the initial bypassing.

DISCUSSION

This work presents a preliminary step towards a principled framework where regression parameters are informed considering spatial hierarchical relationships. This can help to illuminate coherent signals to describe coastal processes at a conceptual level, even from a satellite-derived yearly shoreline dataset. Further examples from a range of sites will be presented. The framework can be extended to nonlinear relationships (e.g., with the use of GAM type models) which are known to be important in coastal processes. Additionally, more advanced techniques exist to capture periodic signals and long memory processes, many of which have been explored in different contexts while modelling coastal response, that could be incorporated into this framework.

REFERENCES

Bishop-Taylor, R., Nanson, R., Sagar, S., & Lymburner, L. (2021). Mapping Australia's dynamic coastline at mean sea level using three decades of Landsat imagery. *Remote Sensing of Environment*, *267*, 112734. https://doi.org/10.1016/j.rse.2021.112734

Kennedy, D. M., McCarroll, R. J., Fellowes, T. E., Gallop, S. L., Pucino, N., McSweeney, S. L., Vila-Concejo, A., Reef, R., Yuan, R., Carvalho, R., Quang, T. H., & Ierodiaconou, D. (2023). Drivers of seasonal and decadal change on an estuarine beach in a fetch-limited temperate embayment. *Marine Geology*, *463*, 107130. https://doi.org/10.1016/j.margeo.2023.107130

Luijendijk, A., Hagenaars, G., Ranasinghe, R., Baart, F., Donchyts, G., & Aarninkhof, S. (2018). The State of the World's Beaches. *Scientific Reports*, *8*(1), 6641. https://doi.org/10.1038/s41598-018-24630-6

Simmons, J. A., Splinter, K. D., & Beuzen, T. (2019). DATA-DRIVEN MODELLING OF SHORELINE EVOLUTION. *Coastal Sediments 2019*, 733-741. https://doi.org/10.1142/9789811204487_0065

Vos, K., Harley, M. D., Splinter, K. D., Simmons, J. A., & Turner, I. L. (2019). Sub-annual to multi-decadal shoreline variability from publicly available satellite imagery. *Coastal Engineering*, *150*, 160-174. https://doi.org/10.1016/j.coastaleng.2019.04.004

Wikle, C. K., Berliner, L. M., & Cressie, N. (1998). Hierarchical Bayesian space-time models. *Environmental and Ecological Statistics*, *5*(2), 117-154. https://doi.org/10.1023/A:1009662704779