**Title**

Using state space time series analysis on wetland bird species to formulate effective bioindicators in the Barberspan wetland.

**Background to research question**

For many years wetlands have been considered useless environments. This was the case locally and globally. The importance of wetlands was not recognised worldwide. In South Africa, back in the 1900s, the government agricultural branch instructed farmers to convert wetland areas into land that is conducive to classical farming. This led to the loss of many wetland regions across South Africa.

Apart from bad farming practices, anthropogenic factors have also led to the loss and/or degradation of wetlands. This is a present reality even after the importance of wetlands has been realised.

Wetlands are important for many things. It acts as a water filtration system for settlements nearby, it also helps with flood control and is a source of water for both humans and animals. This is especially important as South Africa has a generally climate.

There is currently a citizen science dataset (CWAC) that contains bird counts of various wetlands in South Africa. SANBI is currently collaborating with FIAO and SEEC to utilize this dataset to aid wetland conservation decision makers. The number of birds and biodiversity of species in a wetland can be a good indicator to show if a wetland is healthy or not.

**Research Question**

How can the CWAC dataset aid in providing an indicator of wetland health based on the population dynamics of its bird species?

**Research Statement**

The abundance of bird species and the richness of biodiversity amongst bird species over time can give us an indicator to how healthy a wetland is. If there are few birds in a wetland then one can assume the wetland has experienced some sort of change that is driving the bird species away. Given how closely linked an ecosystem is with the species that live there, any change that drives away species is more likely a negative change. Likewise, if there is one or two bird species that dominate a wetland (poor biodiversity) we would have reason to believe the ecosystem is not performing the way it should and thus weakening the health of the wetland.

**Research Hypothesis**

The CWAC dataset will present a pattern of wetland health over time in terms of bird abundance and bird biodiversity. Higher abundance and richer biodiversity would equate to a healthy wetland while low abundance or high abundance and poor biodiversity will equate to an unhealthy wetland ecosystem (need proof for this).

The CWAC data may also prove to be inutile given that the data capturing process is not supervised for each count. It is solely based on confidence that citizens scientists will adhere to the data capturing requirements put in place by CWAC. This could lead to inconsistent practices across wetlands, counters, or time. If there is a high level of inconsistency or high levels of erroneous data capturing amongst the bird counters, we could see no noticeable trends in abundance or biodiversity in the data.

**Research Goals**

Currently the citizen scientist bird count dataset (CWAC) is not being utilised in any way to make decisions regarding wetland conservation. The goal of this study is to analyse the CWAC data and produce effective metrics and visualisations to aid the process in wetland conservation.

**Research Objectives**

1. Perform explanatory data analysis on the CWAC dataset for the Barberspan wetland.
2. Develop a state space time series model to model the bird counts across time (individually and Barberspan as a whole).
3. Formulate biodiversity indices (using hill numbers approach) to see biodiversity of the bird species over time.

Report Sections

* Initial pages
* Introduction
  + Problem description
  + Background to research
  + Purpose of research and objectives (research aim will be here and then broken down to objectives)
* Literature Review
  + Intro to population dynamics.
  + SSMs (use methodological structure applied to state space-models eg. How ssms are used in animal movement, for fisheries stocks, etc, describe frequentist and Bayesian approach and explain why you’ve chosen Bayesian approach)
  + Evaluating those methods. (implicit)
  + Bioindicators (briefly include birds as bioindicators pdf, Shannon, simpson, LPI) – use methodological structure (maybe)
  + Evaluating those bioindicators (implicit)
* Methodology
  + Study area (barberspan, location, climate, water level, surrounding areas/anthropogenic vulnerability, species and why barberspan as pilot (it’s a Ramsar site))
  + Data + EDA (CWAC counting practices)
  + State space model (explain why you’ve chosen ssms and why Bayesian approach)
  + Biodiversity indicators
* Results (add predictions of future values, express least squared errors for the model, show priors you’ve chosen)
* Discussion
* Conclusion and future work