Leveraging Sparse Observations to Predict Species Abundance Across Space and Time

Md Zahidul Islam[1], **Cameron S. Fletcher**[2], **Ke Sun**[3], **Amir Dezfouli**[4], **Iadine Chades**[5]
[1]Khulna University, Khulna 9100
[2]CSIRO Environment, Australian Tropical Science & Innovation Precinct, James Cook University, Townsville, QLD 4811
[3]CSIRO Data61, Eveleigh, NSW 2015
[4]BIMLOGIQ, Sydney, Australia
[5]Environmental Informatics Hub, Monash University, Clayton, VIC 3800
zahid@cse.ku.ac.bd, Cameron.Fletcher@csiro.au, Ke.Sun@data61.csiro.au, amir@bimlogiq.com, iadine.chades@monash.edu

Abstract

Biodiversity is declining globally at an unprecedented rate. Managers urgently need to allocate limited resources to control pest species where interventions have the highest ecological impact. However, many species are hard to detect, and data collection is often expensive, irregular, and incomplete, thus posing significant challenges for machine learning models that traditionally require large and regular datasets. We present a novel deep learning architecture that estimates the spatiotemporal abundance of hard-to-detect species from sparse, zero-inflated, and irregular data. Our method combines Graph Convolutional Networks (GCNs) to model spatial dependencies across monitoring sites with Recurrent Neural Networks (RNNs) to capture long-range temporal dynamics explicitly addresses the challenges of data sparsity, heterogeneity, and irregular sampling. We apply our model to the Crown-of-Thorns Starfish (COTS) on Australia's Great Barrier Reef, a species with devastating impact on coral reefs and a major target of pest control programs. Our method significantly outperforms baseline approaches and the current resource-intensive approach, manta-tow surveillance, in both accuracy and detectability. Simulations indicate a 20% increase in starfish removal efficiency over a year, enabling more effective coral protection. This work demonstrates how tailored deep learning methods can overcome ecological data limitations and substantially improve conservation outcomes.

Code — https://github.com/zahidcseku/STZipN
Datasets — ⟨CodeURL⟩/tree/main/data
Supplementary info. — ⟨CodeURL⟩/tree/main/SI

Introduction

Worldwide, pest invasions threaten biodiversity and agricultural production, accelerating extinction rates and causing a significant economic burden (Diagne et al. 2021; Cuthbert et al. 2021). Between 1970 and 2017, damages increased six-fold each decade to an average of US$18.6 billion per year, while management investments increased by less than two-fold over the same period (Diagne et al. 2021). This widening gap highlights the urgent need for more effective, data-driven pest management strategies. However, when pests are hard to detect, poorly mapped, or spread unpredictably, collecting reliable spatial data is challenging due to cost, logistical difficulties, remote locations and intensive labor requirements. These efforts divert resources away from active management (Britton, Pegg, and Gozlan 2011; Larson et al. 2011). Yet, because rapid action is critical for effective pest control, managers must make decisions about where and when to intervene even when they have limited information (Chadès et al. 2011; Ahmed et al. 2025).

The case of the Crown-of-Thorns Starfish (COTS) control program on Australia's Great Barrier Reef (GBR) exemplifies these challenges. The GBR, a UNESCO World Heritage site and global biodiversity hotspot, is under increasing threat from ocean warming, intensifying cyclones, and outbreaks of COTS (Pratchett et al. 2014). Between 1984 and 2013, COTS were responsible for 42% of all coral lost on the GBR (De'Ath et al. 2012). Of the major threats facing the GBR, they are the most directly manageable driver of coral decline. In response, the Australian government has committed $AUD160M to COTS control efforts between 2018-2030 (Matthews et al. 2024). In practice, COTS are culled at a series of 10 hectare sites distributed around a reef. A COTS control effort, or a voyage, includes *surveillance* and *culling* activities. Sites with the highest estimated COTS densities are prioritized for culling, as removing individuals from these locations minimizes coral loss. To identify these priority sites, managers use a broad-area surveillance method called manta-tow (Miller, Jonker, and Coleman 1964). Although manta-tow has relatively low accuracy, managers consider that it is sufficient for detecting high-density aggregations over large areas (Fernandes et al. 1990). COTS densities fluctuate over time due to culling and natural processes, including the emergence of hidden individuals and movement between sites. Approximately every 50–60 days, manta-tow surveys are conducted to update reef-wide information and reprioritize culling operations. While manta-tows are vital for targeting efforts, they are resource-intensive, diverting effort that could otherwise be invested directly in culling.

Motivated by the need to increase program efficiency by improving COTS abundance estimation and site prioritization (Westcott et al. 2020), we propose a deep learning model called Spatio-Temporal Zero-Inflated Poisson Net-

work (STZipN). STZipN predicts pest abundance at predefined locations based on sparse and irregular historical abundance data. Our approach explicitly tackles three key challenges common in ecological datasets: spatial and temporal sparsity (Liu, Zhao, and Song 2025), zero inflation (Kong et al. 2021), and irregular sampling (Liu, Zhao, and Song 2025; Ahmed et al. 2025). The proposed model includes an RNN-based temporal information encoder and a GCN-based spatial information encoder and is trained with a custom zero-inflated Poisson loss function. We show that STZipN can estimate the spatiotemporal distribution of COTS abundance using historical cull data alone, significantly improving the efficiency of control efforts. We also derive valuable insights into the dispersal of COTS from the learned graph in our model, demonstrating that our approach can be useful to other spatio-temporal ecological problems. Our work makes three key contributions:

- We address a critical pest management problem affecting a globally significant ecosystem, the Great Barrier Reef of Australia, underexplored by the AI community.
- We introduce a novel deep learning model that combines GCNs and RNNs under a zero-inflated Poisson loss to learn from sparse, irregular ecological data.
- Our approach outperforms current best practices and offers strong potential for deployment in collaboration with field managers.

Related Work

Deep learning models can effectively process large temporal datasets enabling opportunities to predict pest abundance in real-time to guide day-to-day decision-making (Capinha et al. 2021; Fulton et al. 2019). However, it is nontrivial to design high performing spatiotemporal deep learning models for pest abundance estimation (Martínez-Minaya et al. 2018; Nielsen et al. 2014; Strebel et al. 2022; Vanhatalo, Hosack, and Sweatman 2017; Deneu et al. 2021). In applied ecology, most deep learning approaches are built on the successes of computer vision and Convolution Neural Networks (CNNs). Applications include coral reef monitoring (Nunes et al. 2020), fish species classification (Siddiqui et al. 2017) and species abundance estimation (Høye et al. 2021; Deneu et al. 2021; Norouzzadeh et al. 2018). However, CNNs usually operate on spatial grid structures that limit the scales at which they can track spatiotemporal dependencies. This is an important limitation for many pest management systems that experience rare long-distance dispersal events and exhibit complex spatiotemporal dynamics (Fletcher and Westcott 2013) that are not suitable to represent as grid structures.

Zero-inflated count data with excess zeros are ubiquitous in ecological datasets (Morley et al. 2018; Martin et al. 2005; Kong et al. 2021). Traditional regression approaches tend to underestimate zeros and overestimate non-zeros in such data, while mixture models like zero-inflated Poisson can better address both the excess zeros and overdispersion. Gaussian process-based approaches have been proposed for zero-inflated data (Hegde, Heinonen, and Kaski 2018; Neelon 2019), offering flexible Bayesian alternatives, though they face computational challenges for large-scale spatiotemporal applications and require careful handling of the zero-inflation mechanism.

In species abundance estimation, data collection sites are not distributed on regular spatial grids and observations are not made at fixed intervals. This spatial and temporal sparsity can be addressed by combining Graph Convolution Networks (GCNs) and Recurrent Neural Nets (RNNs) (Wilson, Tan, and Luo 2018; Jin et al. 2023). GCNs are able to learn spatial patterns from non-grid structures (Kipf and Welling 2017). RNNs are able to embed long-term temporal correlations from irregular time series (Elman 1990). Crucially, these deep learning methods can accommodate zero-inflated observations without requiring explicit mixture modeling or complex likelihood specifications, instead learning appropriate representations directly from data. GCNs and RNNs can infer latent structures from real-world empirical data and leverage the structures to learn complex spatiotemporal dependencies in observations without requiring complex feature engineering, or knowledge of the underlying ecological processes (Zhao et al. 2020; Yan, Xiong, and Lin 2018; Yu, Yin, and Zhu 2018; Wilson, Tan, and Luo 2018).

Our STZipN by design prioritizes simplicity, while more recent architectures including adaptive GNNs such as (Wu et al. 2020; Bai et al. 2020) and diffusion-based GCNs such as (Li et al. 2018) offer promising future improvements. Our use-case presents a novel application of spatial temporal GNNs which have been widely applied to traffic networks (Yu, Yin, and Zhu 2018; Bai et al. 2020; Li et al. 2018; Zhao et al. 2020; Jin et al. 2023) and action recognition (Yan, Xiong, and Lin 2018).

Problem Definition

We formulate the problem as a multi-target Poisson regression task where the objective is to estimate the abundance of a species at a set of m sites represented by $\mathcal{S} := \{(\text{lon}_1, \text{lat}_1), (\text{lon}_2, \text{lat}_2), \cdots (\text{lon}_m, \text{lat}_m)\}$, where lon_i and lat_i denote the longitude and latitude of site i. With a slight abuse of notation, we also use $\mathcal{S}$ to denote a $m \times 2$ matrix of the site locations. We create a training set $(\mathbf{X}, \mathbf{Y})$ consisting of t time steps of observations across m sites, where $\mathbf{X} \in \mathbb{R}^{t \times m \times d}$ is the input features and $\mathbf{Y} \in \mathbb{N}_+^{t \times m}$ is the corresponding abundance counts used as target values. $\mathbb{N}_+$ is the set of positive integers including 0. In this paper, the features are abundance counts and the time difference between observations. The time difference is set to 0 at the start (before any culling) and updated continuously as a running count of days since the previous cull. For an index $i \in [t] := \{1, 2, \ldots, t\}$, $\mathbf{X}_i \in \mathbb{R}^{m \times d}$ is a matrix of observed features at the ith time step, and $\mathbf{Y}_i \in \mathbb{N}_+^m$ is a vector of abundance counts observed at time step $i + 1$.

We assume that species observations at a site are conditionally independent and distributed from a Zero-Inflated Poisson distribution with parameters $\pi \in (0, 1)$ and $\lambda > 0$, where π is the probability of zeros and λ is the expected Poisson count. We use $\hat{\mathbf{Y}} \in \mathbb{R}^m$ to represent the random estimated abundance through the Zero-Inflated Poisson network. If $\mathbf{X}$ is observed until time step t, then $\hat{\mathbf{Y}}$ is the prediction for time step $t + 1$.

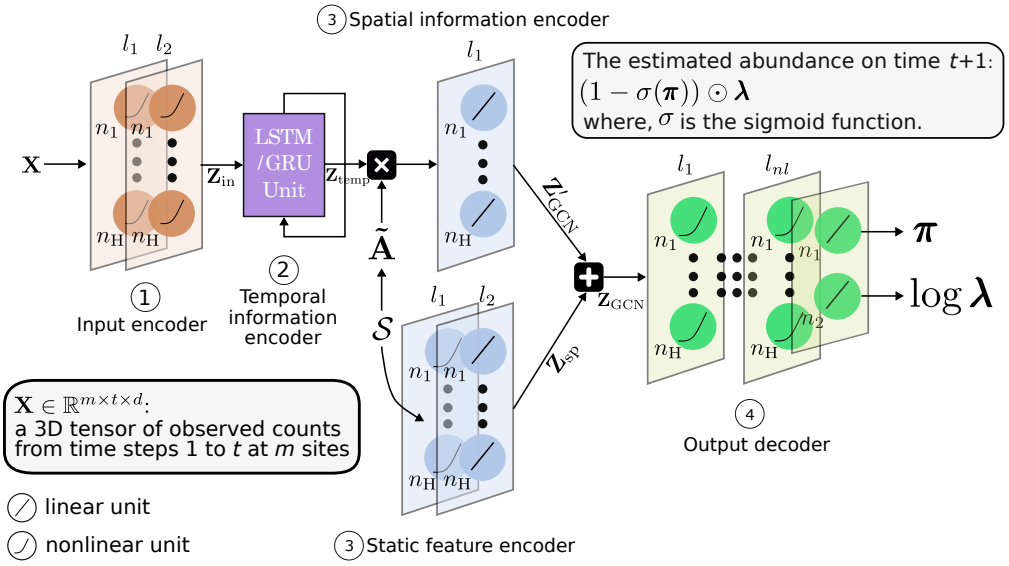

Figure 1: STZipN model architecture. (1) The input encoder transforms the input $\mathbf{X}$ into higher-dimensional latent feature representations. (2) The temporal encoder captures temporal context from these latent feature representations. (3) Spatial information is incorporated through two components: a static feature encoder that encodes site locations, and a spatial encoder that uses a GCN to aggregate temporal features from neighbouring sites. (4) The output decoder combines the spatiotemporal features to estimate the parameters $\boldsymbol{\pi}$ and $\log\boldsymbol{\lambda}$ of the STZipN model, which are used to predict abundance at time $t+1$.

Spatio-Temporal Zero-Inflated Poisson Network

Our proposed Spatio-Temporal Zero-Inflated Poisson Network (STZipN) model has four main modules (Fig. 1).

Input Encoder. The *input encoder* is a Multi-Layer Perceptron (MLP) consisting of two layers of Exponential Linear Units (ELU) (Clevert and Hochreiter 2016), with H units per layer. It encodes the raw input features $\mathbf{X}$ to a higher dimensional latent space (similar to Tran et al. 2021). For simplicity, we use H ELU units for all the hidden layers. The output of the input encoder is given by:

$$\mathbf{Z}_{\text{in}} = \text{ELU}\left(\text{ELU}\left(\mathbf{X}\cdot\mathbf{W}_{\text{in}}^{(1)}+\mathbf{b}_{\text{in}}^{(1)}\right)\cdot\mathbf{W}_{\text{in}}^{(2)}+\mathbf{b}_{\text{in}}^{(2)}\right)$$

where $\mathbf{W}_{\text{in}}^{(1)}\in\mathbb{R}^{d\times\text{H}}$ and $\mathbf{b}_{\text{in}}^{(1)}\in\mathbb{R}^{\text{H}}$ are the weights and biases corresponding to the input to the first hidden layer. Similarly, $\mathbf{W}_{\text{in}}^{(2)}\in\mathbb{R}^{H\times H}$ and $\mathbf{b}_{\text{in}}^{(2)}\in\mathbb{R}^{H}$ are associated with the first hidden layer to the output layer. $\mathbf{Z}_{\text{in}}\in\mathbb{R}^{m\times t\times\text{H}}$ corresponds to the H dimensional embedding of the d dimensional input features.

Temporal Information Encoder. The *temporal information encoder* captures temporal dependencies using a single layer recurrent network that takes $\mathbf{Z}_{\text{in}}$ as input and produces a new temporal representation:

$$\mathbf{Z}_{\text{temp}} = \text{TE}\left(\mathbf{Z}_{\text{in}}\cdot\mathbf{W}_{\text{temp}}+\mathbf{b}_{\text{temp}}\right)$$

where $\mathbf{W}_{\text{temp}}\in\mathbb{R}^{l\cdot\text{H}\times\text{H}}$ and $\mathbf{b}_{\text{temp}}\in\mathbb{R}^{l\cdot\text{H}}$ are the learnable parameters of the temporal encoder (TE). For LSTM, parameters of the TE are $\mathbf{W}_{\text{temp}}\in\mathbb{R}^{4\cdot\text{H}\times\text{H}}$ and $\mathbf{b}_{\text{temp}}\in\mathbb{R}^{4\cdot\text{H}}$ correspond to the input, forget, output and candidate cell gates. Similarly, for GRU, $\mathbf{W}_{\text{temp}}\in\mathbb{R}^{3\cdot\text{H}\times\text{H}}$ and $\mathbf{b}_{\text{temp}}\in\mathbb{R}^{3\cdot\text{H}}$ correspond to the reset, update and candidate cell gates. The output $\mathbf{Z}_{\text{temp}}\in\mathbb{R}^{m\times t\times\text{H}}$ is a temporal encoding of $\mathbf{Z}_{\text{in}}$. To obtain the representation for timestep $t+1$, the hidden state from t is fed back to TE that eventually retains the trends in the observations until $t-1$ to generate representations for timestamp t.

Spatial Information Encoder. The *spatial information encoder* consists of a single-layer GCN and a two-layer MLP. The GCN processes site features by aggregating features from the nearby sites and outputs $\mathbf{Z}'_{\text{GCN}}\in\mathbb{R}^{m\times t\times\text{H}}$. The MLP maps the site locations $\mathcal{S}$ into a latent space $\mathbf{Z}_{\text{sp}}\in\mathbb{R}^{m\times 1\times\text{H}}$. The two encodings $\mathbf{Z}'_{\text{GCN}}$ and $\mathbf{Z}_{\text{sp}}$ are combined and passed through an ELU layer to produce the spatiotemporal feature encoding $\mathbf{Z}_{\text{GCN}}\in\mathbb{R}^{m\times t\times\text{H}}$.

Unique to our problem as well as many species abundance estimation problems is that the observations are not explicitly structured as graphs which is required by the GCN. We construct the graph from the site locations $\mathcal{S}$ following (Yu, Yin, and Zhu 2018) where each site is a vertex and the edges are determined based on the distance between sites D_{ij}. We construct the graph using a kernel width learned form the observation data. We compute the weight $A_{ij}(\tau,\epsilon)$ corresponding to sites i and j in our graph $\mathbf{A}:=(A_{ij})_{m\times m}$ using distance D_{ij}, kernel width τ and cut-off threshold ϵ:

$$A_{ij}(\tau,\epsilon) := \exp\left(-\frac{D_{ij}^2}{\tau^2}\right)\cdot\left[\!\!\left[\exp\left(-\frac{D_{ij}^2}{\tau^2}\right)>\epsilon\right]\!\!\right],$$

where $[\![\textit{expression}]\!]$ is 1 if the *expression* is true, 0 otherwise. The adjacency matrix defines the graph connectivity where $A_{ij}(\tau,\epsilon)>0$. In STZipN, τ is a hyperparameter depending

on how fast the species move. The threshold ϵ governs the sparsity of matrix $\mathbf{A}$. We found that $\epsilon = 10^{-5}$ is a good empirical value and use it throughout. We compute the distance between sites i and j using the Haversine formula (Robusto 1957) as:

$$D_{ij} := 2 \arcsin \sqrt{\begin{array}{l} \sin^2\left(\frac{\text{lat}_i - \text{lat}_j}{2}\right) + \\ \cos(\text{lat}_i)\cos(\text{lat}_j)\sin^2\left(\frac{\text{lon}_i - \text{lon}_j}{2}\right) \end{array}}$$

All pairwise site distances are represented by a matrix $\mathbf{D} := (D_{ij})_{m \times m}$. Note that our graph representations such as $\mathbf{A}(\tau, \epsilon)$ and the corresponding degree matrix $\mathbf{Deg}(\tau, \epsilon)$ depend on τ and ϵ. For simplicity, we denote $\mathbf{A} := \mathbf{A}(\tau, \epsilon)$ and $\mathbf{Deg} := \mathbf{Deg}(\tau, \epsilon)$. We normalize $\mathbf{A}$ as $\tilde{\mathbf{A}} := \mathbf{Deg}^{-\frac{1}{2}}\mathbf{A}'\mathbf{Deg}^{-\frac{1}{2}}$ for numerical stability. $\mathbf{A}' := \mathbf{A} + \mathbf{I}_m$ is the adjacency matrix with self-loops, $\mathbf{Deg} = \text{diag}\{\mathbf{A}' \cdot \mathbf{1}_\text{m}\}$ where $\mathbf{1}_m$ denotes a m dimensional vector of 1 and the $\text{diag}(.)$ function constructs a diagonal matrix from the result of the operation $\mathbf{A}' \cdot \mathbf{1}_m$ (Kipf and Welling 2017).

In our model, the GCN aggregates the temporal encoding of the m sites $\mathbf{Z}_\text{temp}$ using $\tilde{\mathbf{A}}$ and the output of the GCN can be expressed as:

$$\mathbf{Z}_\text{GCN} := \text{ELU}\left(\begin{array}{l} \underbrace{\tilde{\mathbf{A}} \cdot \mathbf{Z}_\text{temp} \cdot \mathbf{W}_\text{GCN}}_{\mathbf{Z}_{\text{GCN}'}} + \mathbf{1}_\text{H} \cdot \mathbf{b} + \\ \underbrace{\text{ELU}(\mathcal{S} \cdot \mathbf{W}^{(1)}_\text{sp} + \mathbf{1}_\text{H} \cdot \mathbf{b}^{(1)\mathsf{T}}_\text{sp}) \cdot \mathbf{W}^{(2)}_\text{sp}}_{\mathbf{Z}_\text{sp}} \end{array}\right),$$

where $\mathbf{W}_\text{GCN} \in \mathbb{R}^{\text{H} \times \text{H}}$ is learnable weights in the GCN submodule. The static encoder is a two-layer perceptron network i.e., $\mathbf{Z}_{sp} = \text{MLP}(\mathcal{S})$. The output of the static encoder is a matrix $\mathbf{Z}_\text{sp} \in \mathbb{R}^{m \times \text{H}}$ representing static features. The shape of $\mathbf{Z}_{\text{GCN}'}$ is $m \times t \times \text{H}$. The parameters of the static enocoder network are $\mathbf{W}^{(1)}_\text{sp} \in \mathbb{R}^{2 \times \text{H}}$, $\mathbf{W}^{(2)}_\text{sp} \in \mathbb{R}^{\text{H} \times \text{H}}$, $\mathbf{b}^{(1)}_\text{sp} \in \mathbb{R}^\text{H}$ and $\mathbf{b}^{(2)}_\text{sp} \in \mathbb{R}^\text{H}$. Bias terms corresponding to the GCN layer and the second layer of the static encoder is represented by $\mathbf{b}$.

Output Decoder. The *output decoder* uses the spatiotemporal encoding $\mathbf{Z}_\text{GCN}$ and estimates two parameters (per site) $\boldsymbol{\pi} \in \mathbb{R}^m$ and $\log \boldsymbol{\lambda} \in \mathbb{R}^m$ of a Zero-Inflated Poisson distribution (Boone, Stewart-Koster, and Kennard 2012) denoted by $\Pr(\mathbf{Y} \mid \boldsymbol{\pi}, \boldsymbol{\lambda})$. For a site i, the prediction for next time step is $\hat{\text{Y}}^i \sim \Pr(\text{Y}^i | \pi^i, \lambda^i)$ where Y^i, π^i, and λ^i are the observation and parameters of the Zero-Inflated Poisson distribution corresponding to site i. $\Pr(\text{Y}^i = 0 | \pi^i, \lambda^i) = \pi^i + (1 - \pi^i)e^{-\lambda^i}$ and $\Pr(\text{Y}^i = y | \pi^i, \lambda^i) = (1 - \pi^i)e^{-\lambda^i}(\lambda^i)^y / y!$ where $y > 0$. The output decoder consists of a MLP network i.e.,

$$[\log \boldsymbol{\lambda}, \boldsymbol{\pi}] := \\ \text{ELU}\left(\text{ELU}\left(\mathbf{Z}_\text{GCN} \cdot \mathbf{W}^{(1)}_\text{out} + \mathbf{b}^{(1)}_\text{out}\right) \cdot \mathbf{W}^{(2)}_\text{out} + \mathbf{b}^{(2)}_\text{out}\right) \\ \cdots \mathbf{W}^{(nl)}_\text{out} + \mathbf{b}^{(nl)}_\text{out}$$

where the number of layers nl is a hyperparameter. The parameters $\mathbf{W}^{(l)}_\text{out} \in \mathbb{R}^{\text{H} \times \text{H}}$ and $\mathbf{b}^{(l)}_\text{out} \in \mathbb{R}^\text{H}$ correspond to layers $l = \{1, 2, \ldots, nl - 1\}$ and parameters $\mathbf{W}^{(nl)}_\text{out} \in \mathbb{R}^{\text{H} \times 2}$ and $\mathbf{b}^{(nl)}_\text{out} \in \mathbb{R}^2$ correspond to the nl^{th} layer. From the STZipN model, the estimated abundance is computed as $\mathbb{E}(\hat{\mathbf{Y}}) = (1 - \sigma(\boldsymbol{\pi})) \odot \boldsymbol{\lambda}$, where σ is the sigmoid function, $\mathbb{E}$ is the expectation w.r.t $\Pr(\mathbf{Y} \mid \boldsymbol{\pi}, \boldsymbol{\lambda})$, and $\odot$ is element-wise product.

Zero inflated Poisson Loss Function

The parameters of deep neural networks are estimated by minimizing a loss function such as the negative log-likelihood. To fit to our problem, we develop the negative log-likelihood of our zero inflated sparse targets including two independent distributions — a Bernoulli distribution to model the binary outcome of a target being zero and a Poisson distribution to model the positive occurrences of the target (Kong et al. 2021; Bai et al. 2025). Consequently, the zero values in the model are generated by the Bernoulli process and also by the Poisson process due to zero occurrences. The log-likelihood $\mathcal{L}$ for zero and nonzero occurrences is:

$$\mathcal{L} := \sum_{i=1}^{m} \left(\begin{array}{l} \sum_{j:\text{Y}^i_j = 0} \log\left(\pi^i + (1 - \pi^i)e^{-\lambda^i}\right) + \\ \sum_{k:\text{Y}^i_k > 0} \log\left((1 - \pi^i)\frac{{\lambda^i}^{\text{Y}^i_k} e^{-\lambda^i}}{\text{Y}^i_k!}\right) \end{array}\right),$$

where $j : \text{Y}^i_j = 0$ and $k : \text{Y}^i_k > 0$ represents the zero and non-zero observations at site i respectively.

Training

STZipN outputs are functions of $\boldsymbol{\Theta}$ (includes all the parameters stated above), $\mathbf{X}$ and $\mathcal{S}$ i.e., $\boldsymbol{\pi}(\boldsymbol{\Theta}, \mathbf{X}, \mathcal{S})$ and $\boldsymbol{\lambda}(\boldsymbol{\Theta}, \mathbf{X}, \mathcal{S})$. We estimate $\boldsymbol{\Theta}$ by minimizing $-\mathcal{L}$ using Adam (Kingma and Ba 2015). We remove sites with missing observations from loss calculations. The hyperparameters of the STZipN model include τ, nl, H and the learning rate. We use a multi-step decaying learning rate policy starting at an initial value, and then decrease the learning rate by 0.2 every 50 epochs. In the search space, we vary τ from 800 to 1300, nl from 1 to 6, H from 16 to 256 and the learning rate from 0.0001 to 0.01 (see supplementary information *Hyperparameter search*). We use early stopping to prevent overfitting.

To preserve the temporal dependencies in our observations, we split the validation and test data from the tail. First, we split the dataset into two parts - for training (75%) and testing (25%). Then, we optimize $\mathcal{L}$ on a validation set of 20% of the training set in our hyperparameter search.

Experimental Results

Experiment Settings

Datasets and evaluation. We assessed the daily estimates of our model based on real-world past COTS culling history across 144 reefs against the cull dataset span over 982 days collated into 190 voyages from November 5, 2018 through July 14, 2021. We use the culled COTS number as a proxy for the true number of COTS at a location, which is not directly observable. Because our data is zero-inflated, standard RMSE and MAE are not appropriate for evaluation. Following (Kong et al. 2021), we adapt the RMSE and MAE to

Model	zMAE	zRMSE
Historical average	37.48	54.77
Per site historical average	34.48	52.93
Last seen	31.63	53.05
ExpertR	40.54	133.03
CatBoost	34.56	77.41
ZINB	42.21	77.07
IDW	42.89	103.08
OK	43.62	88.48
STZipN	**24.50**	**42.73**

Table 1: Comparison of STZipN model predictions with baseline COTS estimation models.

measure the models' performance on zero-inflated data by considering the zero and nonzero outputs separately in our evaluation. We define

$$\text{zRMSE} = \frac{\sum_{i=0}^{N_{\text{test}}} \sqrt{\frac{\alpha}{|I_0|} \sum_{j \in I_0} \left(\hat{y}_j^i\right)^2 + \frac{(1-\alpha)}{|I_{>0}|} \sum_{j \in I_{>0}} \left(y_j^i - \hat{y}_j^i\right)^2}}{N_{\text{test}}},$$

$$\text{zMAE} = \frac{\sum_{i=0}^{N_{\text{test}}} \frac{\alpha \sum_{j \in I_0} \left|\hat{y}_j^i\right|}{|I_0|} + \frac{(1-\alpha) \sum_{j \in I_{>0}} \left|y_j^i - \hat{y}_j^i\right|}{|I_{>0}|}}{N_{\text{test}}},$$

where $y^{(i)} \in \mathbb{R}^{1 \times L}$ is the true observed vector on the i-th timestamp where L is equal to the number of sites, I_0 and $I_{>0}$ represents the zero and non-zero observations in $y^{(i)}$ respectively, and $\hat{y}^{(i)} \in \mathbb{R}^{1 \times L}$ is the predictions of a model. The variable α is a relative importance of the zero part. As both zero and nonzero parts are evaluated separately and weighted by their relative importance (α and $1-\alpha$), the metrics reflect true performance by not favouring the models predicting very small values, i.e. $y_j^{(i)} \approx 0 \ \forall j \in \{1, \cdots, L\}$.

Baselines. Because manta-tow surveillance is designed for large-scale, coarse detection, its low accuracy makes direct comparison with our model predictions uninformative. Instead, we evaluate the prediction error of our model against a suite of baseline abundance estimation approaches including simple historical heuristics (historical average, per site historical average, last seen), expert-guided regression (ExpertR) (S. Fletcher and Westcott 2016), CatBoost (Prokhorenkova et al. 2018), classical statistical method like Zero-Inflated Negative Binomial (ZINB) (Yau, Wang, and Lee 2003) and, geostatistical methods such as Ordinary Kriging (OK) (Cressie 1988) and Inverse Distance Weighting (IDW) (Lu and Wong 2008):

- **Historical average** uses the average number of COTS observed (culled) in the historical periods of recorded data as the estimated number of COTS at the sites.
- **Per site historical average** uses the average number of COTS observed at a site in the historical periods as the estimated number of COTS at the corresponding site.
- **Last seen** the number of COTS observed at the site last time as an estimate of the number of COTS.

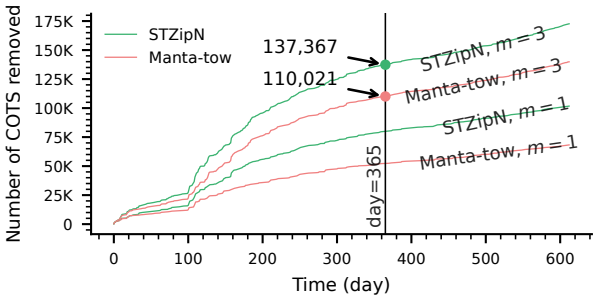

Figure 2: STZipN's recommendations would have resulted in a 20% increase in COTS removed over the first year of the control program compared to manta-tow-guided prioritization. Assuming three sites are managed per day ($m = 3$).

- **ExpertR** estimates COTS by combining the most recent observed density at that site (where available) with a weighted set of correlated factors, selected based on expert research opinion, including year, latitude, mean COTS density for reefs within ± 100 km latitude, mean density for the current reef, and mean densities at the site's immediate neighboring sites on both sides (with neighborhood typically defined by adjacent monitoring points or the spatial scale relevant for COTS movement within reefs).
- **CatBoost** is a high-performance algorithm that constructs ensembles of decision trees to predict continuous numerical outcomes.
- **ZINB** is a standard classical baseline for over dispersed and zero-inflated count data.
- **OK** estimates values at unobserved locations by computing a weighted average of nearby sampled values, where the weights are determined by the spatial autocorrelation captured in a variogram model.
- **IDW** estimates values by averaging the values of nearby sampled values, assigning higher weights to observed locations that are closer to the target location.

For the spatial interpolation methods and CatBoost we used observations from the prior 56 days following (S. Fletcher and Westcott 2016). For OK, we experimented with linear, spherical and Gaussian variogram models.

Results

Table 1 summarizes results comparing the zRMSE and zMAE for the baseline methods on the COTS dataset, highlighting the effectiveness of the STZipN model. Our approach consistently outperforms current practices, a conventional machine learning method, and spatial interpolation techniques in estimating COTS abundance.

While manta-tow surveillance data are a poor predictor of absolute abundance, they are currently used to prioritize sites with the highest estimated COTS densities for culling. Because this prioritization relies on relative rather than absolute abundance, it can tolerate consistent underestimation but is highly sensitive to low precision. We com-

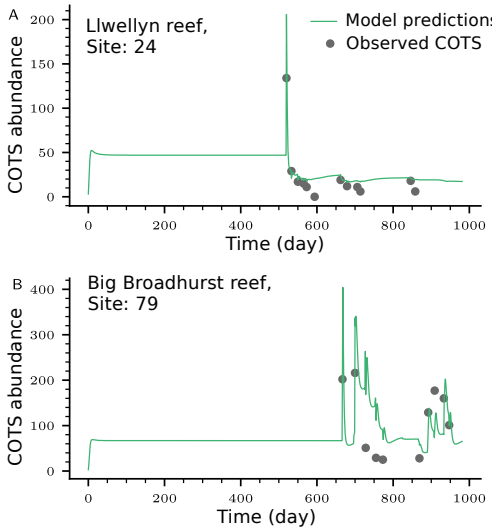

Figure 3: Examples of STZipN's ability to estimate COTS abundance under sparse observations at (**A**) Llwellyn Reef and (**B**) Broadhurst Reef. STZipN predicts COTS abundance on a given day using only culling observations available up to that day. The x-axis represents the timeline of the COTS control program, and the y-axis shows the number of COTS culled.

pared site prioritization based on current practice (manta-tow) with prioritization based on STZipN. Due to its higher precision and more consistent abundance estimates, STZipN recommended the optimal culling sequence more frequently than manta-tow. STZipN's recommendations strongly correlated with ideal site rankings derived *a posteriori* from actual COTS removal over 172 voyages (mean Spearman rank correlation coefficient $\rho = 0.69$, $p = 0.06$). In contrast, manta-tow rankings correlated weakly with actual cull rankings ($\rho = 0.22$, $p = 0.34$; see Supplementary *Evaluation*). Implementing the respective cull ranks, we estimate that 19.91% to 34.56% more COTS could have been removed compared to manta-tow at the end of one year of control, with effectiveness increasing over time (Fig. 2).

Robustness in low observations. Our model is robust in predicting abundance even when the true observations are scarce. We assessed the day-to-day predictions of abundance at a site from our model using only past culling data available up to that day. Fig. 3A shows that our model correctly predicts low COTS at site 24 at the Llwellyn reef on day 520 and beyond, despite having only a single earlier high observation. Fig. 3B shows our model's ability to predict high and low abundance even with no observations prior to day 667.

Manual cross-checking of the data confirms the influence of observations at neighbouring sites in both cases: for site 24 at Llewellyn reef, the neighbouring sites have 44 observations before day 520 and for site 79 at Big Broadhurst reef, the neighbouring sites have 18 observations preceding day 667 (see discussion on the graph construction in the Materials and Methods section). These predicted estimates with no or few observations can be attributed to the GCN's ability to learn feature representations from these neighbouring sites.

We also note that in absence of prior information our model overestimates abundance (e.g., predicting 50 COTS where ground truth was zero). However, it correctly identifies high and low abundance levels which is essential to planning effective control strategies.

Ablation study. To assess the contribution of spatial and temporal components, we performed an ablation study comparing model variants (Table 2). The full STZipN model which uses both the spatial and temporal encoders performs the best among the compared models (zMAE $= 24.50$ COTS/day, zRMSE $= 42.73$ COTS/day and $\mathcal{L} = 35.46$). Removing either encoder degrades performance. Without the spatial encoder (Temporal model), performance declines across all metrics (zMAE = 28.49, zRMSE = 46.35, $\mathcal{L}$ = 39.10). Removing the temporal encoder (Spatial model) further worsens results (zMAE = 31.53, zRMSE = 48.71, $\mathcal{L}$ = 40.86). Excluding temporal and spatial encoders results in the worst performing model (*Trivial* model: zMAE $= 37.30$ COTS/day, zRMSE $= 55.07$ COTS/day and $\mathcal{L} = 49.66$).

Sensitivity analysis. We measured how much the model's output at a given site changes in response to variations in the input observations across space and time. Our model's predictions at a site are mostly sensitive to the input observations collected from sites within 1200 meters and within the preceding 1 to 14 days (Fig. 4). We obtain the sensitivity plot by computing the gradients of the model's output with respect to the inputs because it efficiently captures local response of model outputs to input variations (see *Sensitivity analysis* in the supplementary for more details). This sensitivity to observations within 1200 meters and 14 days aligns with the operational practices of the culling program: a new voyage begins 14 days after the last voyage. Each site is only culled once per voyage, because it is hard to distinguish a live COTS from a recently culled COTS. STZipN leverages data from nearby sites (within 1200m) on both the current voyage (i.e. within a few days) and the most recent previous voyage (i.e. 14 days ago). Fig. 4 also shows that the sensitivity decreases with increasing distance in time and space, confirming that more distant sites and older observations have no to little influence on abundance estimations.

Neighbourhood learning. Learning representations from geographical neighbours using a GCN is a novel contribution to ecological modelling. For the COTS problem, the observations are collected from 1,368 sites. These sites constitute the nodes of the graph that the GCN uses. We construct the edges using distance and a learned spectral kernel width from the cull data. Fig. 5A shows the locations of the observation sites and Fig. 5B shows an example of the graph structure over the Lady Musgrave reef and Fairfax Islands reef. Each site learns feature representations from its connected neighbours. *Predictions at low observation sites* in the supplementary reveals key structural properties of the graph. For example, there are 133 subgraphs (connected components) in the graph that we constructed from data collected from 144 reefs. There are 22 isolated nodes (or sites) in the graph. Because there are fewer connected components than there are reefs, this information suggests that some

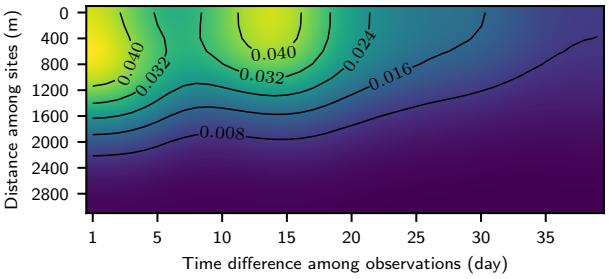

Figure 4: Effects of spatial and temporal information on STZipN predictions. Model predictions are most influenced by observations made within 15 days and from neighbouring sites within 1,200 meters. The x-axis represents the temporal lag (in days) between cull observations, and the y-axis shows the spatial distance (in meters) between sites.

Model	Loss $\mathcal{L}$	zMAE	zRMSE
Trivial	49.66	37.30	55.07
Spatial	40.86	31.53	48.71
Temporal	39.10	28.49	46.35
Spatiotemporal (STZipN)	**35.46**	**24.50**	**42.73**

Table 2: Results of the ablation study. The models Trivial, Spatial and Temporal refer to modifications of our STZipN model. For both zMAE and zRMSE, $\alpha = 0.5$.

sites learn useful patterns from sites located at other reefs. These properties are likely common in ecological monitoring networks, particularly for pests or species inhabiting fragmented habitats with weakly connected metapopulations. In such cases, GCN-based models offer a powerful alternative to traditional regression approaches by capturing complex spatial correlations in sparse and irregular data.

Discussion

STZipN enables improved efficiency of pest management by leveraging deep learning to process sparse spatiotemporal data collected during past control efforts, and using these predictions to guide future control efforts. Applied to Crown-of-Thorns Starfish management on the Great Barrier Reef (Fletcher, Bonin, and Westcott 2020; S. Fletcher and Westcott 2016; Westcott et al. 2021), our model can serve as a substitute for the resource-intensive manta-tow surveillance currently used to guide day-to-day operations. Our approach has superior accuracy and higher detectability than baseline approaches and manta-tow (Fig. 2 and Fig. S5). This is critical because it enables better prioritization of high density sites to deploy control efforts. Our results show that our model recommendations would lead to $\approx$ 20% more COTS removed in one year (Fig. 2B), saving more coral and decreasing the pressure on an already threatened ecosystem.

Once trained, our model can be deployed as a real-time on-site decision-support tool via a web service (Singh 2021). Our model does not require additional resources for data

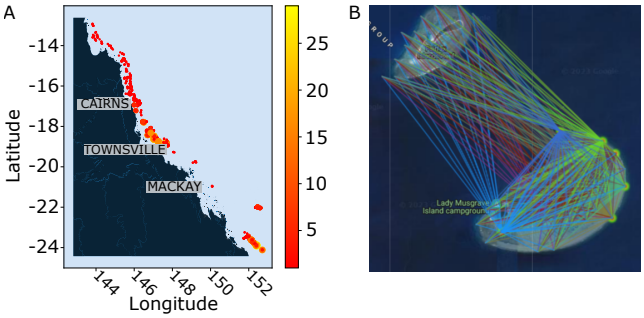

Figure 5: (**A**) Site locations marked by proportion of samples (red gradient) showing the sparsity of observations. (**B**) A part of the graph learned by STZipN model over the Lady Musgrave reef and Fairfax Islands reef showing that a site learns features from other sites within the same reef as well as from sites at nearby reefs. An edge colour represents one individual site's connectivity.

collection and provides instantaneous abundance estimates. Like many deep neural networks, our model is not easily interpretable (Pichler and Hartig 2023). However, our sensitivity analysis and graph structure analysis show how the model correctly learns the spatiotemporal dependencies observed in COTS populations (Fig. 4). Together, these two analyses provide transparency about the model's ecological realism that can increase trust and understanding.

Our method tackles three key challenges in applied ecology — sparse data, zero inflation, and complex spatiotemporal dependencies (Reynolds, Thompson, and Russell 2011; Caley et al. 1996) - by using STZipN, which is built on spatiotemporal modelling principles that capture sparsity, zero inflation, and dependence structure. This structure improves COTS estimates by $\approx$ 13% COTS/day (zMAE, Table 2). While we are still in the early stages of understanding the superior performance of deep learning models in these challenging contexts, STZipN is naturally designed based on core principles of spatiotemporal modeling.

Beyond COTS, our approach is generalizable to other sparsely monitored species including koalas (Dissanayake et al. 2019) and migratory birds (Munson et al. 2009). While the model architecture is reusable, the graph and learned parameters must be adapted to reflect species-specific dispersal and habitat use (Yates et al. 2018). Future work could explore transfer learning to support rapid adaptation across ecological domains (Pan and Yang 2010).

Our findings contribute a generalizable framework for species management under data scarcity, especially relevant as ecological management costs continue to outpace available resources (Diagne et al. 2021). Our model offers a scalable, low-cost solution for large-scale, spatially dispersed species. More broadly, we demonstrate that deep learning can be effective in conservation even without "big data" (Chadès and Nicol 2016), highlighting a path forward for AI to support urgent biodiversity protection goals (Tuia et al. 2022; Capinha et al. 2021; Elhamod et al. 2023; Reynolds et al. 2025).

Acknowledgments

We thank Cheng Soon Ong (CSIRO) and Dan Pagendam (CSIRO) for providing useful feedback on this manuscript.

References

Ahmed, N.; Roth, M.; Hallman, T. A.; Robinson, W. D.; and Hutchinson, R. A. 2025. Spatial clustering of citizen science data improves downstream species distribution models. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, 27775–27783.

Bai, L.; Yao, L.; Li, C.; Wang, X.; Wang, C.; and Gao, J. 2020. Adaptive Graph Convolutional Recurrent Network for Traffic Forecasting. In *Proceedings of the 34th Conference on Neural Information Processing Systems (NeurIPS)*.

Bai, S.; Ji, Y.; Liu, Y.; Zhang, X.; Zheng, X.; and Zeng, D. D. 2025. Alleviating Performance Disparity in Adversarial Spatiotemporal Graph Learning Under Zero-Inflated Distribution. *Proceedings of the AAAI Conference on Artificial Intelligence*, 39: 11436–11444.

Boone, E.; Stewart-Koster, B.; and Kennard, M. 2012. A hierarchical zero-inflated Poisson regression model for stream fish distribution and abundance. *Environmetrics*, 23(3): 207–218.

Britton, J. R.; Pegg, J.; and Gozlan, R. E. 2011. Quantifying imperfect detection in an invasive pest fish and the implications for conservation management. *Biological Conservation*, 144(9): 2177–2181.

Caley, M. J.; Carr, M. H.; Hixon, M. A.; Hughes, T. P.; Jones, G. P.; and Menge, B. A. 1996. Recruitment and the local dynamics of open marine populations. *Annual Review of Ecology and Systematics*, 27(1): 477–500.

Capinha, C.; Ceia-Hasse, A.; Kramer, A. M.; and Meijer, C. 2021. Deep learning for supervised classification of temporal data in ecology. *Ecological Informatics*, 61: 101252.

Chadès, I.; Martin, T. G.; Nicol, S.; Burgman, M. A.; Possingham, H. P.; and Buckley, Y. M. 2011. General rules for managing and surveying networks of pests, diseases, and endangered species. *Proceedings of the National Academy of Sciences*, 108(20): 8323–8328.

Chadès, I.; and Nicol, S. 2016. Small data call for big ideas. *Nature*, 539(7627): 31–31.

Clevert, T., Djork-Arnéand Unterthiner; and Hochreiter, S. 2016. Fast and Accurate Deep Network Learning by Exponential Linear Units (ELUs). In *International Conference on Learning Representation*.

Cressie, N. 1988. Spatial prediction and ordinary kriging. *Mathematical geology*, 20(4): 405–421.

Cuthbert, R. N.; Pattison, Z.; Taylor, N. G.; Verbrugge, L.; Diagne, C.; Ahmed, D. A.; Leroy, B.; Angulo, E.; Briski, E.; Capinha, C.; Catford, J. A.; Dalu, T.; Essl, F.; Gozlan, R. E.; Haubrock, P. J.; Kourantidou, M.; Kramer, A. M.; Renault, D.; Wasserman, R. J.; and Courchamp, F. 2021. Global economic costs of aquatic invasive alien species. *Science of The Total Environment*, 775: 145238.

Deneu, B.; Servajean, M.; Bonnet, P.; Botella, C.; Munoz, F.; and Joly, A. 2021. Convolutional neural networks improve species distribution modelling by capturing the spatial structure of the environment. *PLOS Computational Biology*, 17(4): 1–21.

De'Ath, G.; Fabricius, K. E.; Sweatman, H.; and Puotinen, M. 2012. The 27–year decline of coral cover on the Great Barrier Reef and its causes. *Proceedings of the National Academy of Sciences*, 109(44): 17995–17999.

Diagne, C.; Leroy, B.; Vaissière, A. C.; E. G., R.; Roiz, D.; Jarić, I.; Salles, J. M.; Bradshaw, C. J. A.; and F., C. 2021. High and rising economic costs of biological invasions worldwide. *Nature*, 592(6): 571–576.

Dissanayake, R. B.; Stevenson, M.; Allavena, R.; and Henning, J. 2019. The value of long-term citizen science data for monitoring koala populations. *Scientific reports*, 9(1): 10037.

Elhamod, M.; Khurana, M.; Manogaran, H. B.; Uyeda, J. C.; Balk, M. A.; Dahdul, W.; Bakis, Y.; Bart, H. L.; Mabee, P. M.; Lapp, H.; Balhoff, J. P.; Charpentier, C.; Carlyn, D.; Chao, W.-L.; Stewart, C. V.; Rubenstein, D. I.; Berger-Wolf, T.; and Karpatne, A. 2023. Discovering Novel Biological Traits From Images Using Phylogeny-Guided Neural Networks. In *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, KDD '23, 3966–3978.

Elman, J. L. 1990. Finding Structure in Time. *Cognitive Science*, 14(2): 179–211.

Fernandes, L.; Marsh, H.; Moran, P.; and Sinclair, D. 1990. Bias in manta tow surveys of Acanthaster planci. *Coral Reefs*, 9: 155–160.

Fletcher, C. S.; Bonin, M. C.; and Westcott, D. A. 2020. An ecologically-based operational strategy for COTS Control:Integrated decision making from the site to the regional scale. *Reef and Rainforest Research Centre Limited, Cairns*.

Fletcher, C. S.; and Westcott, D. A. 2013. Dispersal and the design of effective management strategies for plant invasions: matching scales for success. *Ecological Applications*, 23(8): 1881–1892.

Fulton, E. A.; Blanchard, J. L.; Melbourne-Thomas, J.; Plaganyi, E. E.; and Tulloch, V. J. D. 2019. Where the Ecological Gaps Remain, a Modelers' Perspective. *Frontiers in Ecology and Evolution*, 7.

Hegde, P.; Heinonen, M.; and Kaski, S. 2018. Variational zero-inflated Gaussian processes with sparse kernels. *arXiv preprint arXiv:1803.05036*.

Høye, T. T.; Ärje, J.; Bjerge, K.; Hansen, O. L. P.; Iosifidis, A.; Leese, F.; Mann, H. M. R.; Meissner, K.; Melvad, C.; and Raitoharju, J. 2021. Deep learning and computer vision will transform entomology. *Proceedings of the National Academy of Sciences*, 118(2): e2002545117.

Jin, G.; Liu, L.; Li, F.; and Huang, J. 2023. Spatio-Temporal Graph Neural Point Process for Traffic Congestion Event Prediction. *Proceedings of the AAAI Conference on Artificial Intelligence*, 37: 14268–14276.

Kingma, D. P.; and Ba, J. 2015. Adam: A Method for Stochastic Optimization. In *International Conference on Learning Representations*.

Kipf, T. N.; and Welling, M. 2017. Semi-Supervised Classification with Graph Convolutional Networks. In *International Conference on Learning Representations*.

Kong, S.; Bai, J.; Lee, J. H.; Chen, D.; Allyn, A.; Stuart, M.; Pinsky, M.; Mills, K.; and Gomes, C. P. 2021. Deep Hurdle Networks for Zero-Inflated Multi-Target Regression: Application to Multiple Species Abundance Estimation. In *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence*, IJCAI'20.

Larson, D. L.; Phillips-Mao, L.; Quiram, G.; Sharpe, L.; Stark, R.; Sugita, S.; and Weiler, A. 2011. A framework for sustainable invasive species management: Environmental, social, and economic objectives. *Journal of Environmental Management*, 92(1): 14–22.

Li, Y.; Yu, R.; Shahabi, C.; and Liu, Y. 2018. Diffusion Convolutional Recurrent Neural Network: Data-Driven Traffic Forecasting. In *International Conference on Learning Representations (ICLR '18)*.

Liu, Z.; Zhao, X.; and Song, Y. 2025. RDPI: A Refine Diffusion Probability Generation Method for Spatiotemporal Data Imputation. *Proceedings of the AAAI Conference on Artificial Intelligence*, 39: 12255–12263.

Lu, G. Y.; and Wong, D. W. 2008. An adaptive inverse-distance weighting spatial interpolation technique. *Computers & Geosciences*, 34(9): 1044–1055.

Martin, T. G.; Wintle, B. A.; Rhodes, J. R.; Kuhnert, P. M.; Field, S. A.; Low-Choy, S. J.; Tyre, A. J.; and Possingham, H. P. 2005. Zero tolerance ecology: improving ecological inference by modelling the source of zero observations. *Ecology Letters*, 8(11): 1235–1246.

Martínez-Minaya, J.; Cameletti, M.; Conesa, D.; and Pennino, M. G. 2018. Species distribution modeling: a statistical review with focus in spatio-temporal issues. *Stochastic environmental research and risk assessment*, 32: 3227–3244.

Matthews, S. A.; Williamson, D. H.; Beeden, R.; Emslie, M. J.; Abom, R. T. M.; Beard, D.; Bonin, M.; Bray, P.; Campili, A. R.; Ceccarelli, D. M.; Fernandes, L.; Fletcher, C. S.; Godoy, D.; Hemingson, C. R.; Jonker, M. J.; Lang, B. J.; Morris, S.; Mosquera, E.; Phillips, G. L.; Sinclair-Taylor, T. H.; Taylor, S.; Tracey, D.; Wilmes, J. C.; and Quincey, R. 2024. Protecting Great Barrier Reef resilience through effective management of crown-of-thorns starfish outbreaks. *PLOS ONE*, 19(4): 1–29.

Miller, I.; Jonker, M.; and Coleman, G. 1964. Crown-of-thorns starfish and coral surveys using the manta tow technique. *Australian Institute of Marine Science, Townsville, Australia*, 155: 8–14.

Morley, J. W.; Selden, R. L.; Latour, R. J.; Frölicher, T. L.; Seagraves, R. J.; and Pinsky, M. L. 2018. Projecting shifts in thermal habitat for 686 species on the North American continental shelf. *PLOS ONE*, 13(5): 1–28.

Munson, M. A.; Webb, K.; Sheldon, D.; Fink, D.; Hochachka, W. M.; Iliff, M.; Riedewald, M.; Sorokina, D.; Sullivan, B.; Wood, C.; et al. 2009. The ebird reference dataset, version 1.0. *Cornell Lab of Ornithology and National Audubon Society, Ithaca, NY*.

Neelon, B. 2019. Bayesian zero-inflated negative binomial regression based on Pólya-Gamma mixtures. *Bayesian Analysis*, 14(3): 829.

Nielsen, J. R.; Kristensen, K.; Lewy, P.; and Bastardie, F. 2014. A Statistical Model for Estimation of Fish Density Including Correlation in Size, Space, Time and between Species from Research Survey Data. *PLOS ONE*, 9(6): 1–15.

Norouzzadeh, M. S.; Nguyen, A.; Kosmala, M.; Swanson, A.; Palmer, M. S.; Packer, C.; and Clune, J. 2018. Automatically identifying, counting, and describing wild animals in camera-trap images with deep learning. *Proceedings of the National Academy of Sciences*, 115(25): E5716–E5725.

Nunes, J. A. C.; Cruz, I. C.; Nunes, A.; and Pinheiro, H. T. 2020. Speeding up coral reef conservation with AI-aided automated image analysis. *Nature Machine Intelligence*, 2(6): 292–292.

Pan, S. J.; and Yang, Q. 2010. A Survey on Transfer Learning. *IEEE Transactions on Knowledge and Data Engineering*, 22: 1345–1359.

Pichler, M.; and Hartig, F. 2023. Machine learning and deep learning — A review for ecologists. *Methods in Ecology and Evolution*, 14(4): 994–1016.

Pratchett, M. S.; Caballes, C. F.; Rivera-Posada, J. A.; and Sweatman, H. P. 2014. Limits to understanding and managing outbreaks of crown-of-thorns starfish (Acanthaster spp). *Oceanography and Marine Biology: An Annual Review*, 52: 133–200.

Prokhorenkova, L.; Gusev, G.; Vorobev, A.; Dorogush, A. V.; and Gulin, A. 2018. CatBoost: unbiased boosting with categorical features. In *Proceedings of the 32nd International Conference on Neural Information Processing Systems*, NIPS'18, 6639–6649.

Reynolds, J. H.; Thompson, W. L.; and Russell, B. 2011. Planning for success: Identifying effective and efficient survey designs for monitoring. *Biological Conservation*, 144(5): 1278–1284.

Reynolds, S. A.; Beery, S.; Burgess, N.; Burgman, M.; Butchart, S. H.; Cooke, S. J.; Coomes, D.; Danielsen, F.; Di Minin, E.; Durán, A. P.; Gassert, F.; Hinsley, A.; Jaffer, S.; Jones, J. P.; Li, B. V.; Mac Aodha, O.; Madhavapeddy, A.; O'Donnell, S. A.; Oxbury, W. M.; Peck, L.; Pettorelli, N.; Rodríguez, J. P.; Shuckburgh, E.; Strassburg, B.; Yamashita, H.; Miao, Z.; and Sutherland, W. J. 2025. The potential for AI to revolutionize conservation: a horizon scan. *Trends in Ecology & Evolution*, 40(2): 191–207.

Robusto, C. C. 1957. The Cosine-Haversine Formula. *The American Mathematical Monthly*, 64(1): 38–40.

S. Fletcher, C.; and Westcott, D. A. 2016. Strategies for Surveillance and Control: Using Crown-of-Thorns Starfish management program data to optimally distribute management resources between surveillance and control. *Reef and Rainforest Research Centre Limited, Cairns*, 1.

Siddiqui, S. A.; Salman, A.; Malik, M. I.; Shafait, F.; Mian, A.; Shortis, M. R.; and Harvey, E. S. 2017. Automatic fish species classification in underwater videos: exploiting pre-trained deep neural network models to compensate for limited labelled data. *ICES Journal of Marine Science*, 75(1): 374–389.

Singh, P. 2021. Deploy Machine Learning Models to Production. *Cham, Switzerland: Springer*.

Strebel, N.; Kéry, M.; Guélat, J.; and Sattler, T. 2022. Spatiotemporal modelling of abundance from multiple data sources in an integrated spatial distribution model. *Journal of Biogeography*, 49(3): 563–575.

Tran, A.; Mathews, A.; Ong, C. S.; and Xie, L. 2021. Radflow: A Recurrent, Aggregated, and Decomposable Model for Networks of Time Series. In *Proceedings of the Web Conference 2021*, WWW '21, 730–742. Association for Computing Machinery.

Tuia, D.; Kellenberger, B.; Beery, S.; Costelloe, B. R.; Zuffi, S.; Risse, B.; Mathis, A.; Mathis, M. W.; van Langevelde, F.; Burghardt, T.; et al. 2022. Perspectives in machine learning for wildlife conservation. *Nature communications*, 13(1): 792.

Vanhatalo, J.; Hosack, G. R.; and Sweatman, H. 2017. Spatiotemporal modelling of crown-of-thorns starfish outbreaks on the Great Barrier Reef to inform control strategies. *Journal of Applied Ecology*, 54(1): 188–197.

Westcott, D. A.; Fletcher, C. S.; Gladish, D.; and Babcock, R. 2021. Monitoring and Surveillance for the Expanded Crown-of-Thorns Starfish Management Program. *Reef and Rainforest Research Centre Limited, Cairns*, 1.

Westcott, D. A.; S. Fletcher, C.; J. Kroon, F.; Babcock, R. C.; Plagányi, E. E.; Pratchett, M. S.; and Bonin, M. C. 2020. Relative efficacy of three approaches to mitigate Crown-of-Thorns Starfish outbreaks on Australia's Great Barrier Reef. *Scientific Reports*, 10(12594).

Wilson, T.; Tan, P.-N.; and Luo, L. 2018. A Low Rank Weighted Graph Convolutional Approach to Weather Prediction. In *2018 IEEE International Conference on Data Mining (ICDM)*, 627–636.

Wu, Z.; Pan, S.; Long, G.; Jiang, J.; Chang, X.; and Zhang, C. 2020. Connecting the Dots: Multivariate Time Series Forecasting with Graph Neural Networks. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*.

Yan, S.; Xiong, Y.; and Lin, D. 2018. Spatial temporal graph convolutional networks for skeleton-based action recognition. In *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence*, AAAI'18/IAAI'18/EAAI'18. AAAI Press.

Yates, K. L.; Bouchet, P. J.; Caley, M. J.; Mengersen, K.; Randin, C. F.; Parnell, S.; Fielding, A. H.; Bamford, A. J.; Ban, S.; Barbosa, A. M.; et al. 2018. Outstanding challenges in the transferability of ecological models. *Trends in ecology & evolution*, 33(10): 790–802.

Yau, K. K. W.; Wang, K.; and Lee, A. H. 2003. Zero-Inflated Negative Binomial Mixed Regression Modeling of Over-Dispersed Count Data with Extra Zeros. *Biometrical Journal*, 45(4): 437–452.

Yu, B.; Yin, H.; and Zhu, Z. 2018. Spatio-temporal graph convolutional networks: a deep learning framework for traffic forecasting. In *Proceedings of the 27th International Joint Conference on Artificial Intelligence*, 3634–3640.

Zhao, L.; Song, Y.; Zhang, C.; Liu, Y.; Wang, P.; Lin, T.; Deng, M.; and Li, H. 2020. T-GCN: A Temporal Graph Convolutional Network for Traffic Prediction. *IEEE Transactions on Intelligent Transportation Systems*, 21(9): 3848–3858.