



Available online at <http://scik.org>

Commun. Math. Biol. Neurosci. 2020, 2020:66

<https://doi.org/10.28919/cmbn/4943>

ISSN: 2052-2541

AN INVESTIGATION OF POWER LAW DISTRIBUTION IN WILDEBEEST (*CONNOCHAETES TAURINUS*) HERDS IN SERENGETI NATIONAL PARK, TANZANIA

LINUS N. KISOMA^{1,*}, COLIN TORNEY², DMITRY KUZNETSOV¹, ANNA C. TREYDTE^{3,4}

¹School of Computational and Communication Science and Engineering, Nelson Mandela African Institution of Science and Technology, Arusha Tanzania

²Department of Mathematics and Statistics, University of Glasgow, UK

³School of Life Science and Bioengineering, Nelson Mandela African Institution of Science and Technology, P. O. Box 447, Arusha Tanzania

⁴Agroecology, Hans-Ruthenberg Institute, University of Hohenheim, Stuttgart, Germany

Copyright © 2020 the author(s). This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract: Animal group dynamics have often been studied by biologists through the use of mathematical models and statistical analyses. Wildebeest herds (*Connochaetes taurinus*) occur in large numbers and follow certain migration patterns throughout the year. However, it is not known whether the aggregation patterns of migrating wildebeest herds follow predictable statistical distributions. In this work, we investigated whether social interactions between individual wildebeest can generate the observed distribution patterns of herds based on empirical data of wildebeest in the Serengeti, Tanzania. We quantified the distribution of real herds by analyzing the frequency distribution of wildebeest counts in aerial survey images collected in 2015. We then used a Lagrangian model of

*Corresponding author

E-mail address: linuskisoma@gmail.com

Received August 12, 2020

animal interactions to simulate individual movement and herd aggregation patterns. We equipped the model with parameter values that matched empirical distributions. Our results from the empirical data analysis reveal that wildebeest herds follow a truncated power law in their aggregation patterns. We claim that this behaviour can be explained by social interactions between individual wildebeest.

Keywords: agent based model; aggregation patterns; group size; frequency distribution; truncated power law.

2010 AMS Subject Classification: 92B05, 92D40, 92D50.

1. INTRODUCTION

Animals are often found in groups such as fish schools, bird flocks, insect swarms and ungulate herds [1-2]. Being in a group helps members to engage in different behavioral activities such as in foraging, predator avoidance, resistance to toxic environmental conditions, reproduction or socialization [3-4]. Foraging animals usually travel in groups and make movement decisions that depend on forage availability but also on their social interactions [1]. The stability and direction of the group depends on the knowledge about quality and location of the food source and the ability of the informed individuals to influence group decisions to move to a desired direction [5].

Wildebeest (*Connochaetes taurinus*) and other ungulates migrate from Serengeti National Park in northern Tanzania to the Masai Mara region in Kenya creating yearly mass movement of millions of individuals of several species moving together in groups [6]. The great migration involves an estimated 1.3 million wildebeest, 200,000 zebra, and a multitude of gazelles, among various other hoofed species [7]. This migration is caused by the search for green and nutritious pasture by wildebeest in the Serengeti plains following rains and freshly green vegetation [6].

The Serengeti wildebeest movement patterns have extensively been studied for decades. Some mathematical models exist in literature describing the effect of competition, predation, and harvesting in a single or multi-species system [8-10]; but little is known about mathematical models and statistical distributions that describe aggregation patterns of wildebeest herds to predict collective behavior resulting from such aggregations.

Collective animal behaviour has been studied using different mathematical models such as Aoki, [11], Reynolds, [12] and Huth & Wissel, [13]. In 2002, Couzin, et al [14] introduced a three dimensional Lagrangian model to study collective behaviour of fish schools and bird flocks. This model considers three types of social interactions between individuals in their relevant neighborhoods that can trigger various movement responses, i.e., repulsion, alignment, and attraction [14]. Couzin's model reveals that different group level behaviour may be caused by minor changes in individual level interactions within a herd. While this principle holds for some fish schools, we wanted to know if it can produce statistical distributions that match empirical data in ungulate herds of wildebeest of the Serengeti. Understanding a specific mathematical model and the statistical distribution of group sizes in wildebeest herds shows their evolutionary fitness as the optimal balance between costs and benefits of an individual wildebeest will determine whether it joins the group [4]. In this study, we analysed long-tailed distributions to predict aggregation patterns in wildebeest herds. This is because long-tailed group size distributions are similar to the aggregation of physical particles [15]. We considered and analysed three heavy tailed-distributions; power law, truncated power law and exponential distributions. We used the python package [16] to identify the best fitting statistical distribution to explain the observed aggregation in wildebeest herds detected in aerial survey images in the Serengeti, Tanzania, in 2015.

The spatial probability distribution of a given species is an important element for understanding the aggregation patterns in that species [15]. Wildebeest in the Serengeti ecosystem aggregate in herds whose sizes are not well defined, i.e., ranging from tens of individuals to up to about 400,000 [17]. Therefore, describing the group distribution with a model will be helpful for predicting movement of wildebeest aggregation. As this species is a key-stone species in the Serengeti ecosystem [18] upon which many other species depend, the results will be useful for further management of the Serengeti ecosystem and its forage resources for the benefit of both residents and migrating species. Further, long-tailed group size distributions have a cutoff size (maximum size) because the population is finite [4]. The cutoff size depends on ecological

conditions like food availability and presence of predators [4] and will be an important contribution to understanding habitat use and species aggregations in the Serengeti.

The aim of our study was to investigate whether social interactions between individual wildebeest can generate the empirical power law distributions of wildebeest aggregations. Specifically, 1) we quantified the distribution of wildebeest herds by analyzing the frequency distribution of empirical wildebeest counts in aerial survey images; 2) we used an agent based model to understand herd aggregations and to investigate whether individual movement parameter values can generate simulated herd distributions that statistically match empirical distributions.

2. MATERIALS & METHODS

2.1 Description of data

Data used in this study are based on aerial photo counts collected from the Serengeti ecosystem between April and May, 2015 showing the distribution of migratory herds of wildebeest [19]. The wildebeest migrations occur in a cycle between the Serengeti in Tanzania and the Masai Mara in Kenya; with most of the movements taking place in Tanzania [18], [20] The area surveyed includes the eastern and southern plains of Serengeti National Park, Loliondo Game Controlled Area, Maswa Game Reserve Area, and Ngorongoro Conservation Area. The images of migratory wildebeest were collected after every 10 seconds by a camera placed at the floor of the aircraft at the start of each transect [19]. Places where wildebeest were not detected returned a count of 0. The photos collected may not contain a complete herd density in each return as the maximum herd size was determined by the image area. The flight altitude was set at 213 m above the ground and the aircraft speed was maintained at 185 km/h. Reconnaissance flights over the surveyed area covered a total straight line distance of 2040 km.

2.2 Empirical data analysis

We analysed the empirical data to identify the presence of power laws. Power laws are probability distributions which take the form $p(x) \propto x^{-\alpha}$, where $p(x)$ is the probability of

obtaining a herd of size x , and α is an exponent or a scaling parameter and lies in the range of $2 < \alpha < 3$ [16]. The scaling parameter α has been found to express the role environmental factors such as temperature, resource availability, and population density on group size distributions [4], [15]. In recent years, several statistical methods for evaluating power law fit have been proposed [4], [15]. In this paper, we used a power law package for identifying the presence of long-tailed distributions that describe our data [16]. We selected this method because it is easy to implement and contains a variety of probability distributions for analysis.

We performed goodness of fit tests of different distributions (power law, truncated power law and exponential distribution). The aim of performing the goodness of fit tests to the photo counts was to decide the distribution that best explains our data [16]. Data were fit to the three distributions in python software (power law python package as described by [16]). We used Kolmogorov-Smirnov tests to test for individual fit and the likelihood ratio R . We used p and R values to compare two distributions at a time and determined the performance of the two fits. We set the p – value threshold to be $p \leq 0.05$ [16]. When comparing two candidate distributions (say D1 and D2), if the likelihood ratio R , between two candidate distributions is positive, then the data are more likely to be in the first distribution (D1) and negative if the data are more likely in the second distribution (D2) [16].

2.3 Comparison of group size distributions from the aerial photo counts

We performed the goodness of fit test of a power law and exponential distributions. We found that $R = 2.292$ and $p = 0.022$. Since R was positive and $p < 0.05$, the power law distribution was chosen as the better fit compared to the exponential distribution. The goodness of fit test of a truncated power law and exponential distributions was performed and we obtained $R = 2.754$ and $p = 0.006$. In fact, R is positive and $p < 0.05$. Therefore, the truncated power law distribution was chosen to be better than the exponential distribution. In the third case, the goodness of fit of a power law and truncated power law distributions was performed. The goodness of fit of these distributions fits was compared and found to be $R = -1.349$ and $p = 0.1$. Therefore, in this case we moderately support the truncated power law distribution as a

better fit (this is because the significance value $p > 0.05$). Hence we conclude that truncated power law distribution gives the best goodness of fit of all the long-tailed distributions considered (see figure1).

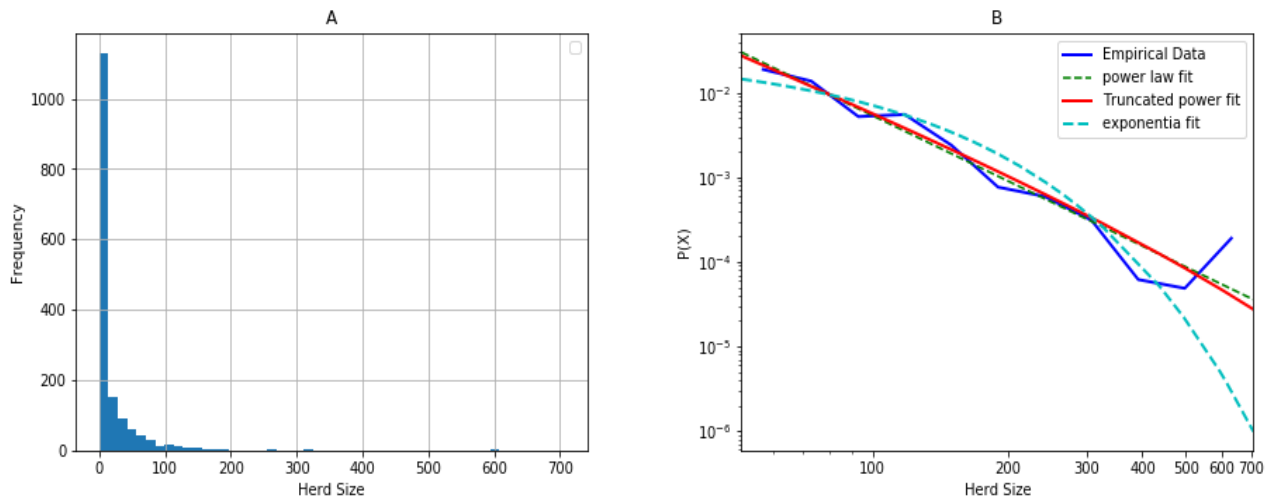


Figure 1: Histogram of the frequency distribution (A) and the power law plot of the probability density function (B) of the aerial photo counts for wildebeest herd sizes.

Most of the herds counted had sizes up to 100 individuals while few herds had sizes beyond 100 individuals to about 700 individuals (Fig. 1A). We observed the presence of a long-tail (B) suggesting the presence of aggregation patterns in wildebeest herds (Fig. 1B). In particular, the truncated power law was observed as the best fit to explain such aggregation patterns.

From the empirical data analysis, we obtained the scaling parameter was $\alpha = 2.561$ and the standard deviation parameter was $\sigma = 0.118$.

2.4 Simulated data

We simulated the Lagrangian model of animal interaction to show individual wildebeest movement and eventually herd aggregation patterns of wildebeest individuals. The resulting distributions and parameters generated from the simulations were compared with the results observed from empirical data analysis.

The model suggests that the collective behavior of wildebeest interactions results from three behavioral movement rules that are exhibited by wildebeest individuals; 1) wildebeest move into

the same direction as their neighbors; 2) they remain close to their neighbors, and; 3) they avoid collisions, i.e., getting too close to each other [5]. These rules were modeled using three distinct contributions to the inter-individual interactions:

- (i) attraction, which ensures no animal remains isolated
- (ii) alignment of velocities, which makes neighboring animals move in the same direction;
- (iii) repulsion (short range repulsion), which prevents proximity that could lead to collisions.

These three types of movement behavior have been used to study a wide range of taxonomic groups, including insects, birds and fish [7].

We further extended the model presented by Couzin [14] into two dimensions, i.e., we investigated how groups of wildebeest merge and split by considering how individuals change their relative positions. Following Couzin [14], if groups are composed of N individuals ($i = 1, 2, 3, \dots, N$) with each individual having a position vector $c_i(t)$ and direction vector $v_i(t)$ moving with speed s_i has a minimum distance α between itself i and others j at all time. The desired direction of travel will then be

$$d_i(t + \Delta t) = - \frac{\sum_{j \neq i}^{n_r} c_{ij}(t)}{\left| \sum_{j \neq i}^{n_r} c_{ij}(t) \right|} \quad (1)$$

Where, $c_{ij}(t) = \frac{c_j(t) - c_i(t)}{|c_j(t) - c_i(t)|}$ is the unit vector in the direction of neighbor j and n_r is the number of animals in the region of repulsion.

If no neighbors are within the zone of repulsion $n_r = 0$, the individual responds to other animals within the zone of alignment (orientation) and the zone of attraction.

The zone of alignment contains n_o detectable neighbors with $r_r \leq |c_j - c_i| < r_o$. The width of this zone is defined as $\Delta r_o = r_o - r_r$

Furthermore, an individual can align itself with other members within the zone of orientation. The preferred travel direction resulting from the zone of orientation is the average of the neighbor's velocities [14]. This zone can be explained by the equation

$$d_i(t + \Delta t) = \frac{\sum_{j \neq i}^{n_o} v_j(t)}{\left| \sum_{j \neq i}^{n_o} v_j(t) \right|} \quad (2)$$

According to Couzin et al., 2002, wildebeest individuals tend to be attracted towards other individuals (to avoid being isolated). The attraction represents the tendency of wildebeest to join groups and to avoid being on the periphery [2]. The zone of attraction contains n_a detectable neighbors with $r_o \leq |c_j - c_i| < r_a$. The width of this zone is defined as $\Delta r_a = r_a - r_o$.

Attraction can be explained as

$$d_i(t + \Delta t) = \frac{\sum_{j \neq i}^{n_a} c_{ij}(t)}{\left| \sum_{j \neq i}^{n_a} c_{ij}(t) \right|} \quad (3)$$

Re writing equations (1 to 3) we have the resulting direction of all group embers to be

$$d'_i(t + \Delta t) = \frac{d_i(t + \Delta t)}{|d_i(t + \Delta t)|} \quad (4)$$

3. MAIN RESULTS

3.1 Agent Based Model on movement patterns and aggregations

To analyse the individual movement and herd aggregations of wildebeest individuals, we investigated the effect of different parameters in the model (Table 1). The values of the parameters were obtained after running the agent based model several times to obtain values that properly describe the movement patterns in the interaction zones (zone of attraction, zone of alignment and zone of repulsion).

| Parameter | Symbol | Unit | Values used |
|-----------------------|------------|------------------|-------------|
| Number of individuals | N | None | 100 – 5,000 |
| Region of repulsion | r_r | units | 0.01 |
| Region of orientation | r_o | units | 0.05 |
| Region of attraction | r_a | units | 6 |
| Speed | s | Units per second | 0.01 |
| Time step | Δt | seconds | 0.01 |

Table 1: Parameters used in the agent based model that were based on the previous model simulation

Individuals in the model start with random orientation and random positions in the region in which each individual can detect at least one neighbor [2]. We simulated the movement behaviour of individuals using parameters like number of individuals N , which represents the assumed population size and ranged between 100 and 5,000 individuals. In each population size, we ran the simulations up to 10,000 time steps, which were assumed to be dynamically stable. The collective behavior of the model was observed by the resulting distribution after each simulation. After 10,000 time steps, we recorded the resulting herds as optimal, and we combined the final herd sizes from different simulations to make a total of 1,600 herds (each simulation gives 100 herds). We analysed the collective behaviour in different interaction zones by varying the parameter values (see table 1). The radius of the interaction zones was tested for different values and we obtained the best values of interaction zones (regions) which are region of repulsion $r_r = 0.01$ units, region of attraction $r_a = 6$ units, region of orientation (alignment) $r_o = 0.05$ units and the time step $\Delta t = 0.01$ seconds (figure 2).

3.2 Comparison of group size distributions from the agent based model

We performed the goodness of fit tests to the agent based model data to decide which distribution best explains the agent based model.

First we performed the goodness of fit test of a power law and exponential distributions and found that $R = 1.351$ and $p = 0.176$. The value of R was positive and $p > 0.05$, here the power distribution had a better fit than the exponential distribution. Second, the goodness of fit test of a truncated power law and exponential distribution was performed and we obtained $R = 2.836$ and $p = 0.004$. As R is positive and $p < 0.05$, the truncated power law distribution fitted the data better than the exponential distribution. In the third case, the goodness of fit of a power law and truncated power law distribution was performed and compared. As $R = -2.291$ and $p = 0.003$, the truncated power law distribution fitted better than power law. Hence we conclude that truncated power law distribution gives the best goodness of fit of all the long-tailed distributions considered (figure 2).

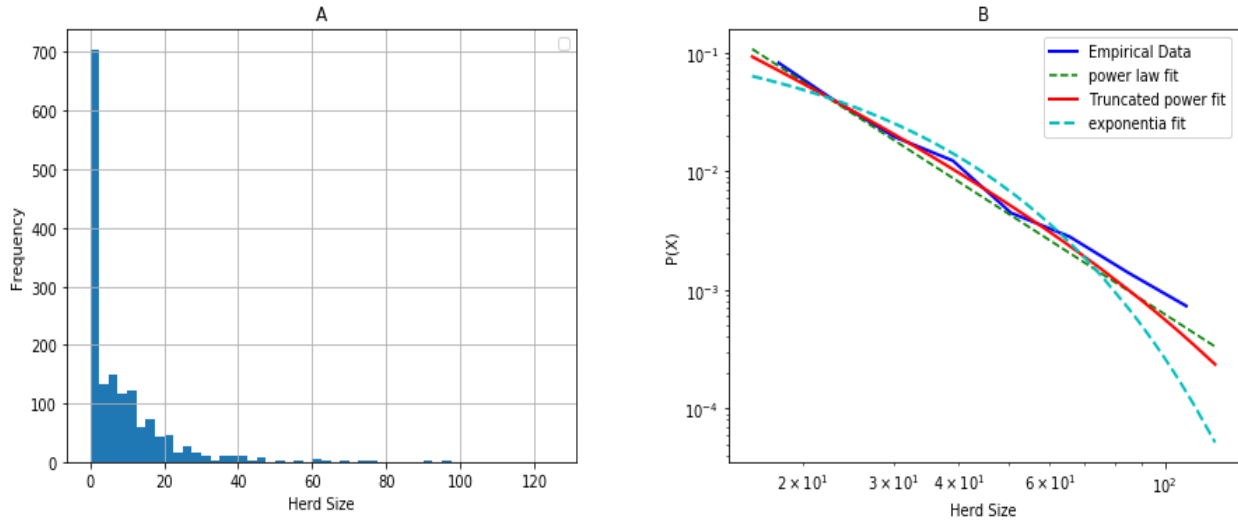


Figure 2: Histogram of the frequency distribution (A) and the power law plot of the probability density function (B) of the agent based model (Lagrangian model) fitting empirical data on wildebeest herd distribution.

In the simulation of distribution of herds of wildebeest from the agent based model most of the herds counted had sizes up to 40 individuals while few herds had sizes beyond 40 individuals to about 120 individuals (Fig. 2A). The resulting smaller group sizes were due to small population size selected for a computer to handle (100-5,000). We observed the presence of a long-tail (Fig. 2B), suggesting the presence of aggregation patterns in wildebeest herd where the truncated power law was observed as the best fit to explain such aggregation patterns.

From the agent based model simulations, we obtained the scaling exponent parameter was $\alpha = 2.809$ and the standard deviation parameter was $\sigma = 0.105$.

4. DISCUSSION

The simulations of our agent-based model exhibited characteristic aggregation patterns that were most similar to the empirical data. We observed a close match between parameters from the empirical data and agent based model. These parameters include the scaling parameter from the power law (α) and the standard deviation σ .

Though there is a difference in the maximum herd size of the aerial photo counts from that of the

agent based model, the aggregation pattern remains to be the similar (Fig.1 and Fig.2). In the empirical data, the maximum herd size was about 700 individuals while in the agent-based model this was about 120 individuals. This shows that in the Serengeti ecosystem wildebeest herd sizes vary from small groups composed of tens of individuals to hundreds of thousands. The similarity in aggregation patterns of wildebeest herds and agent based model is evidence that the selected agent based model generates group size distributions that are long-tailed. This observation relates with earlier studies like [4], where the elementary model of animal aggregation was found consistent with empirical data.

The three forces of attraction, alignment, and repulsion assume that individuals change their orientation in response to the orientation of, at least some, of their neighbors [14]. This was also observed when running our simulations of the agent based model, in which individuals were allowed to detect and join neighbors to form herds. Hence our results confirm that coordination and large scale patterns of wildebeest herds of the Serengeti can be formed from the actions of individuals [2].

Furthermore, the statistical distribution in our results converged to a truncated power law, which matched the empirical data distribution of wildebeest herds. Biologically, this is important because individuals join groups to respond to internal needs (e.g., hunger, thirst, exhaustion) and external needs (e.g., detection and avoidance of predators) [2], [14], [15].

We also identified variations in wildebeest group sizes from the empirical data as well as in the agent based model. This variation signifies that wildebeest respond to different circumstances for their survival. First, depending on the environmental conditions, the size and stability of wildebeest groups may vary. For instance, wildebeest tend to travel longer distances in large groups during the wet season i.e., when food and water are plentiful [20]. Second, the presence of predators causes wildebeest to form larger and stable groups as a mechanism of defense [3]. Third, wildebeest congregate to form larger groups before they navigate across dangerous terrains such as rivers with crocodiles [18]. Other factors that influence group dynamics include habitat open and closed habitats [4]. Habitat openness increases the probability of more

conspecifics to join the group (increased attraction) as open habitat increases the opportunity of wildebeest to see each other and join groups [4]. In contrast, closed habitat leads to smaller and unstable groups because patches of closed habitat prevent individuals from seeing each other [4].

5. CONCLUSIONS

Our agent based model and the empirical data display a truncated power law. These results are in agreement with earlier studies like [4] and [15], where a truncated power law was found to best describe the aggregation patterns in animal (ungulate) herds. In our study we observed that the truncated power law can present a reliable tool for describing aggregation patterns in wildebeest herds of the Serengeti ecosystem. This consistency in empirical data and the model of animal aggregation is evidence that social interaction behaviors alone can lead to large-scale patterns observed in the empirical data. Since wildebeest is a principal species in the Serengeti ecosystem upon which many other species depend, understanding its aggregation patterns will be useful for further management of the Serengeti ecosystem and its forage resources. Although more data is needed from other migrating species like zebra (*Equus burchelli*) and Thomson's Gazelles (*Eudorcas thomsonii*) to quantify the collective aggregation patterns, we believe that this model applies to a wide scale of cases where group size can be large and aggregation is based on minor changes in individual level interactions.

ACKNOWLEDGEMENT

This work was supported by the National Geographic Society through grant number EC-405R-18.

CONFLICT OF INTERESTS

The authors declare that there is no conflict of interests.

REFERENCES

- [1] D. Fortin, M.-E. Fortin, H.L. Beyer, T. Duchesne, S. Courant, K. Dancose, Group-size-mediated habitat selection and group fusion–fission dynamics of bison under predation risk, *Ecology*. 90 (2009), 2480–2490.
- [2] S. Gueron, S.A. Levin, D.I. Rubenstein. The dynamics of herds: From individuals to aggregations. *J. Theor. Biol.* 182 (1996), 85–98.
- [3] Q. Ma, A. Johansson, D.J.T. Sumpter, A first principles derivation of animal group size distributions, *J. Theor. Biol.* 283 (2011), 35–43.
- [4] E. Bonabeau, L. Dagorn, P. Freon, Scaling in animal group-size distributions, *Proc. Nat. Acad. Sci.* 96 (1999), 4472–4477.
- [5] I.D. Couzin, J. Krauset, N.R. Franks, S.A. Levin, Effective leadership and decision-making in animal groups on the move. *Nature*, 433 (2005), 513–516.
- [6] R.M. Holdo, J.M. Fryxell, A.R.E. Sinclair, A. Dobson, R.D. Holt, Predicted Impact of Barriers to Migration on the Serengeti Wildebeest Population, *PLoS ONE*. 6 (2011), e16370.
- [7] S. Grange, P. Duncan, J.-M. Gaillard, A.R.E. Sinclair, P.J.P. Gogan, C. Packer, H. Hofer, M. East, What limits the Serengeti zebra population? *Oecologia*. 140 (2004), 523–532.
- [8] D. Ikanda, C. Packer, Ritual vs. retaliatory killing of African lions in the Ngorongoro Conservation Area, Tanzania. *Endang. Spec. Res.* 6 (2008), 67–74.
- [9] T.H. Fay, J.C. Greeff, Lion, wildebeest and zebra: A predator-prey model. *Ecol. Model.* 196 (2006), 237–244.
- [10] T.D. Sagamiko, N. Shaban, C.L. Nahonyo, O.D. Makinde, Optimal Control of a Threatened Wildebeest-Lion Prey-Predator System in the Serengeti Ecosystem. *Appl. Comput. Math.* 4 (2015), 296–312.
- [11] I. Aoki, A simulation study on the schooling mechanisms in Fish. *Bull. Japan. Soc. Sci. Fisher.* 48 (1982), 1081–1088.
- [12] C.W. Reynolds, Flocks, herds and schools: a distributed behavioral model. *Comput. Graph.* 21(1987), 25–34.
- [13] A. Huth, C. Wissel. The simulation of the movement of fish schools. *J. Theor. Biol.* 156 (1992), 365–385.
- [14] I.D. Couzin, J. Krauset, R. James, G.D. Ruxton, N.R. Franks, Collective Memory and Spatial Sorting in Animal Groups. *J. Theor. Biol.* 218 (2002), 1–11.

- [15] M. Sjöberg, B. Albrechtsen, J. Hjalten, Truncated Power Law: A tool for understanding aggregation patterns in animals? *Ecol. Lett.* 3 (2000), 90–94.
- [16] J. Alstott, E. Bullmore, D. Plenz, powerlaw: A Python Package for Analysis of Heavy-Tailed Distributions, *PLoS ONE*. 9 (2014), e85777.
- [17] C.J. Torney, J.G.C. Hopcraft, T.A. Morrison, I.D. Couzin, S.A. Levin, From single steps to mass migration: the problem of scale in the movement ecology of the Serengeti wildebeest, *Phil. Trans. R. Soc. B.* 373 (2018), 20170012.
- [18] S.A.R. Mduma, Serengeti wildebeest population dynamics: Regulation, limitation and implications for harvesting. PhD thesis, University of British Columbia, 1996.
- [19] C.J. Torney, D.J. Lloyd - Jones, M. Chevallier, D.C. Moyer, H.T. Maliti, M. Mwita, E.M. Kohi, G.C. Hopcraft, A comparison of deep learning and citizen science techniques for counting wildlife in aerial survey images, *Meth. Ecol. Evol.* 10 (2019), 779 – 787.
- [20] J.G.C. Hopcraft, Ecological implications of food and predation risk for herbivores in the Serengeti. Groningen: s.n., 2010, 278 p.