Characterization of Animal Movement Patterns using Information Theory: a Primer

Kehinde Owoeye [1*], Mirco Musolesi [2], Stephen Hailes [1],

1 Department of Computer Science, University College London, London, United Kingdom
2 Department of Geography, University College London, London, United Kingdom

*K.Owoeye@cs.ucl.ac.uk

Abstract

Understanding the movement patterns of animals across different spatio-temporal scales, conditions, habitats and contexts is becoming increasingly important for addressing a series of questions in animal behaviour studies, such as mapping migration routes, evaluating resource use, modelling epidemic spreading in a population, developing strategies for animal conservation as well as understanding several emerging patterns related to feeding, growth and reproduction. In recent times, information theory has been successfully applied in several fields of science, in particular for understanding the dynamics of complex systems and characterizing adaptive social systems, such as dynamics of entities as individuals and as part of groups.

In this paper, we describe a series of non-parametric information-theoretic measures that can be used to derive new insights about animal behaviour with a specific focus on movement patterns, namely Shannon entropy, Mutual information, Kullback-Leibler divergence and Kolmogorov complexity. In particular, we believe that the metrics presented in this paper can be used to formulate new hypotheses that can be verified potentially through a set of different observations and be complementary to existing techniques. We show how these measures can be used to characterize the movement patterns of several animals across different habitats and scales. Specifically, we show the effectiveness in using Shannon entropy to characterize the movement of sheep with Batten disease, mutual information to measure association in pigeons, Kullback-Leibler divergence to study the flights of Turkey vulture, and Kolmogorov complexity to find similarities in the movement patterns of animals across different scales and habitats. Finally, we discuss the limitations of these methods and we outline the challenges in this research area.

Introduction

Information theory has always played an important role in biology [1, 2]. It is a field that is devoted to studying the storage, communication and quantification of information founded by Claude E. Shannon in his influential paper [3] and lies at the interface of mathematics, statistics, computer science and electrical engineering. While initial research in this field was mainly theoretical, we have witnessed a plethora of practical applications in the past decades. For example, concepts and techniques from this field have been used in several fields such as neurobiology [4], pattern recognition [5], cryptology [6], bioinformatics [7], quantum computing [8] and complex systems [9, 10] with significant success.

Recently, due to technological advances, low cost miniaturized sensors have been increasingly adopted for tracking the behaviour of animals across different scales and habitats. These sensors include, but are not limited to, Global Positioning System (GPS) receivers, accelerometers and radio-frequency identification (RFID) tags. This has led to an explosion in the deployment of these sensors in different habitats, across scales by animal behaviour researchers. A natural consequence of this development is that it is now possible to try to quantify and understand a variety of aspects related to animal behaviour such as migration patterns and routes, feeding, reproduction and mating patterns, conservation, monitoring of endangered species, epidemic spreading, resource use, social behaviour and association. In the field of animal behaviour, GPS, accelerometers and cameras are the predominant sensors deployed to measure several behavioural properties of animals. We limit our discussion here to GPS sensors due to their ease of deployment and the apparent simplicity of interpreting the data they produce as well as their consequent popularity relative to other sensors. For example, GPS sensors have been used to study selfish herd behaviour of sheep under threat [11], the hierarchical structures of group dynamics in flocks of pigeons [12], migration patterns in vultures [13], productivity in cows [14], and social relationships in birds [15] just to name a few.

We believe that information-theoretic approaches can provide *complementary* insights in the study of animal behaviour. In other words, these approaches do not replace the existing ones, but they are able to provide additional information about animal behaviour patterns, especially in terms of movement, which are not apparent using other types of analysis. Probabilistic approaches are also usually more robust in presence of noise, a common feature of sensor data. The number of applications of concepts and techniques from information theory to analysis of animal movement in the literature is limited. However, information-theoretic metrics have been used in the past for example to study information flow in animal-robot interactions [16] as well as predator-prey relationships [17, 18] in animals.

In this paper, we discuss how four classic information theoretic metrics, namely Shannon entropy [19], mutual information [20], Kullback-Leibler divergence [21] and Kolmogorov complexity (normalized compression distance) [22] can be effectively applied to the study of animal movement and provide additional *complementary* insights about animal behaviour. In other words, we explore how they can be used as tools for studying animal movement data. Indeed, the goal of this work is not to introduce new metrics, but to demonstrate the potential in using information theoretic concepts to understand animal behaviour. We introduce each metric separately and then we discuss how each metric can be applied to a practical problem, by discussing a case study in detail. More specifically, we demonstrate how these metrics can be used to characterize the movement patterns of animals across different scales and habitats. It is worth noting that these methods do not provide new ground-truth information, but they allow for identifying emergent patterns and formulating hypotheses that can be verified for example by means of further experimental observations in the field. The case studies are mostly based on datasets from the Movebank database [23].

The rest of the paper is organized as follows. In Section 1, we introduce the Shannon entropy as a method for the analysis of characteristics of movement and we demonstrate how it can be used to characterize the movement patterns of sheep with neurodegenerative disease. In Section 2, we describe the notion of mutual information and show how it can be used to measure association as well as reconstruct the flight dynamics in pigeons. We describe the Kullback-Leibler divergence in Section 3 and show how it can be used to characterize the annual movement patterns of the Turkey Vulture. In Section 4, we describe the Kolmogorov complexity (normalized compression distance) and demonstrate how it can be applied to extract the relationships in animal

movement patterns across scales and habitats. We conclude our work by highlighting the challenges and a summary of the methods described in this work.

1 Shannon Entropy

1.1 Overview

The information content of a random variable is defined by the Shannon entropy [19] as a measure that quantifies the level of uncertainty embedded in such variable. In the case of animal movement, Shannon entropy provides a useful quantification of the level of regularity and predictability of the movement of an animal.

More formally, it can be defined as the uncertainty associated to a random variable X with realization x, which can be described by the following equation:

$$H(x) = -\sum_{x} p(x) \ \log \ p(x). \tag{1}$$

where $p(x)$ is the probability density function and the summation is taken over all possible realizations of X. The base of the logarithm is not important and can take any value, provided the same base is used throughout the analysis.

Most of the information theoretic measures that exist today, some of which we will discuss below, are derived from Shannon entropy. However, there are other several definitions of entropy such as Rènyi entropy [24] and Tsallis entropy [25]. The discussion of these metrics is outside the scope of the present article.

Due to the noisy nature of most datasets, probabilistic metrics are becoming increasingly useful for modelling not only animal movement data but in general real world datasets to account for any form of uncertainty inherent in datasets of this nature such as missing data. Therefore, entropy can be used for assessing the overall welfare and well-being of animals instead of a metric like distance travelled. Most animals are known to have a regular activity-rest pattern except under highly unfavourable conditions or when they have some sort of impairment in their general well-being. This implies that animals are supposed to have a relatively high entropy except for the period of harsh conditions when they are either aestivating or hibernating. For this reason, entropy can be used to characterize the movement patterns of animals so as to assess the state of their health. In addition, it can also be used in lieu of tortuosity to describe how tortuous an animal's path is using the turn angle as the input. The conditional entropy is an important element of the conditional mutual information and can be used for example in understanding swarm behaviour [26].

In the following subsection, we will consider a case study illustrating a possible application of Shannon entropy to the study of animal behaviour and, more specifically, to the characterization of the movement patterns of sheep with neurodegenerative diseases.

1.2 Case study: Shannon entropy as a tool for characterizing movement patterns of sheep with neurodegenerative disease

The detection of abnormal locomotion patterns is essential for the early diagnosis of a number of neurodegenerative diseases such as Batten disease in animals. Sheep with neurodegenerative diseases such as Batten disease are known to exhibit repetitive behaviours [27] over time due to gradual loss of motor skills [28] and social awareness [29]. Here, we use Shannon entropy to characterize the movement patterns of a flock of sheep comprising sheep with a natural mutation for Batten disease and their

age matched control (mean age of 2 years) group using the dataset of [30]. We use the trajectory of each sheep sampled every second and compute the distance covered every ten minutes over eleven hours (20:00-7:00) each day for a total period of six days. This time window is chosen in order to minimize the influence of external environmental noise in the dataset. We further bin the resulting distance calculated in order to assign to each sub-intervals symbols. To bin the data, given that we are working with skewed distributions, we use the head/tail classification rule introduced in [31] resulting in 12 bins. First we evaluate this approach on synthetic data where we create two synthetic sheep sampling their distance covered from some random generalized Pareto distributions. We create ten of these distributions with different parameters. The abnormal sheep is allowed to sample from just two of these distributions while the normal sheep is allowed to sample from all. Details of this can be found in S1 Appendix. We compute the entropy and results Fig 1A) show the normal sheep have a higher entropy compared to the abnormal sheep. We further compute the entropy for each sheep (see Table 1 in S1 Appendix) as well as the mean entropy (Fig 1B) for the two groups of sheep. Our results show that the Batten sheep on the average have a lower entropy than the normal sheep with p-values (ANOVA) (0.0076, 0.1042, 0.2628, 0.0065, 0.0234, 0.0205) across the six days respectively. The potential impact of uncontrollable environmental variables such as unfavourable weather conditions is significant and may influence the behaviour of the sheep especially because the experiment was carried out in an open field. Therefore, the result should be interpreted with caution. We compare the entropy of the two groups of sheep with their respective average distance covered in (Fig 1C) and its mean variance (Fig 1D). The Batten sheep can be seen to have covered, on average, a longer distance over the period of observation.

2 Mutual Information

2.1 Overview

We now describe another classic information theoretic measure intimately linked to entropy, called mutual information [20]. The mutual information of two random variables X and Y defines the mutual influence one variable has over the other. Specifically, it quantifies the amount of information in one variable embedded in the other. For this reason, mutual information can be used as a measure of association or social-grouping, for example in the characterization of leader-follower relationship, group coordination and, more generally, collective behaviour [32]. This can be further used to derive a social network [33] and its complementary to the gambit of the group approach [34]. Mutual information can be used to measure non-linear relationships between two variables. More formally, the mutual information of two discrete random variables X and Y, with realizations x and y respectively, is given by:

$$I(X;Y) = \sum_y \sum_x p(x,y) \ \log \ \frac{p(x,y)}{p(x)p(y)}. \tag{2}$$

In this, $p(x)$ and $p(y)$ are the marginal probability distribution functions of X and Y respectively and $p(x,y)$ is the joint probability distribution function of X and Y.

In the case of continuous random variables we have:

$$I(X;Y) = \int_Y \int_X p(x,y) \ \log \ \frac{p(x,y)}{p(x)p(y)}. \tag{3}$$

We now consider a potential application of mutual information, in its application to the study of association in pigeons.

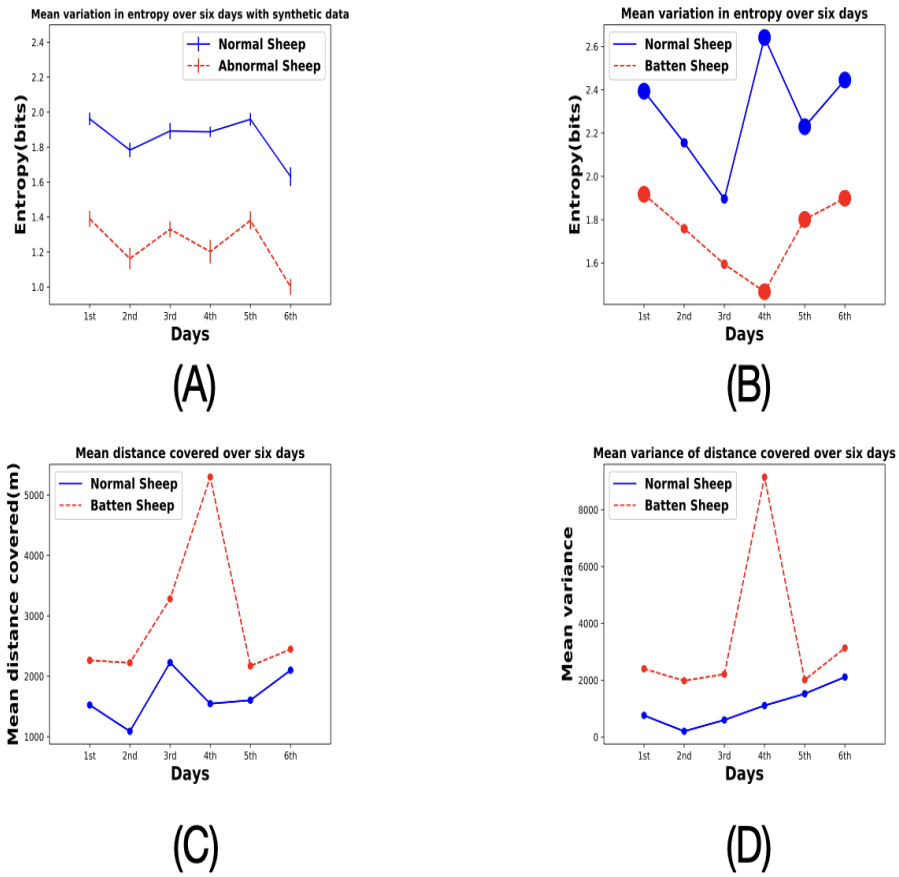

Fig 1. Entropy, distance and variance comparison (A) Mean entropy with synthetic data; (B) Mean entropy of the two groups of sheep across 6 days where the larger circles represent days when the mean difference in entropy are statistically significant (One way ANOVA with $p \leq 0.05$). The sheep affected by Batten disease can be seen to have a lower entropy due to the tendency to repeat the same behaviour over a long period of time; (C) Mean distance covered by the two groups of sheep; (D) Corresponding variance of the mean distance covered.

2.2 Case study: Mutual information for measuring association and leadership in pigeons

As mentioned earlier, the de-facto method used in the animal behaviour community for measuring association is the gambit of the group otherwise known as co-location [34]. To detect significant associations and minimize co-location by chance, a permutation test (random shuffling of associations) is often carried out [35]. However, it is difficult to identify a general method for defining co-location. Also, the directional correlation delay time method (measures how long it takes for one bird to change direction relative to another) used by [12] in reconstructing pigeon flight network structure can only

detect linear relationships leaving the non-linear relationships undetected. Previously in [36], the authors used transfer entropy (measured directed transfer of information) to infer leadership in Zebrafish. We state here that its our aim to detect association using a bidirectional graph and not a directed graph where information flow and its direction is of utmost importance. In addition, transfer entropy may be of limited use in instances where the agents are constantly changing positions relative to one another [37]. For a review of other methods that have been used for measuring leadership and influence, please refer to [38].

We demonstrate how Mutual Information can be used to overcome the limitations associated with the methods above by using it to measure association between pigeons in flight. We use the dataset of [37, 39] and select flight 8 as the result was discussed in the literature in detail. We compute the time-series of the turn angle of each bird followed by the pairwise mutual information of these time series of the nine birds involved in the flight to obtain a distance matrix (see Table 2 in S1 Appendix). As expected, there will always be a certain degree of association between all the birds in the flight, we use a randomization test to determine a threshold for significant pairwise mutual information (S1 Appendix). We further build a social network (Fig 2A) to visualize the flight formation. Our result is consistent with two previous studies on pigeon flight. First, we observe that pigeons do show a hierarchical formation when in flight as seen in (Fig 2A). This is a result consistent with the observations in [12]. Also, we were also able to detect the leader as node M during the flight which is the node M with one edge (Fig 2A)[1] (see video of ground truth trajectory). This result is also consistent with the ground truth in the literature [37]. We compare this approach with three other methods: correlation coefficient, transfer entropy and Granger causality (Fig 2). While it is not straightforward to compare the performance of the four methods, one basis for comparison concerns leadership. Mutual information shows the best performance in terms of leader identification accuracy followed by the correlation coefficient. It is not possible to identify the leader using the transfer entropy and Granger causality. We attribute the poor performance of the transfer entropy and Granger causality to the continuous change in positions of the birds when flying.

3 Kullback-Leibler Divergence

3.1 Overview

Kullback-Leibler divergence (KLD) also known as relative entropy measures the distance between two probability distributions. In animal movement modelling, KLD can be used to quantify changes in behaviour of an individual animal or discrepancies in behaviour in a group. For discrete probability distributions P and Q, KLD is defined as follows:

$$\begin{aligned} D_{Kullback-Leibler}(P||Q) &= \sum P(x) \ \log \ \frac{P(x)}{Q(x)} \\ &= \sum P(x) \ \log \ P(x) - \sum P(x) \ \log \ Q(x) \\ &= H(P) - H(P,Q). \end{aligned} \tag{4}$$

where $H(P,Q)$ is the joint entropy between P and Q and $H(P)$ the entropy of P. For continuous probability distributions P and Q, KLD is given by:

$$D_{Kullback-Leibler}(P||Q) = \int_{-\infty}^{+\infty} P(x) \ \log \ \frac{P(x)}{Q(x)} dx. \tag{5}$$

[1] https://www.youtube.com/watch?v=19srQkTVTGE

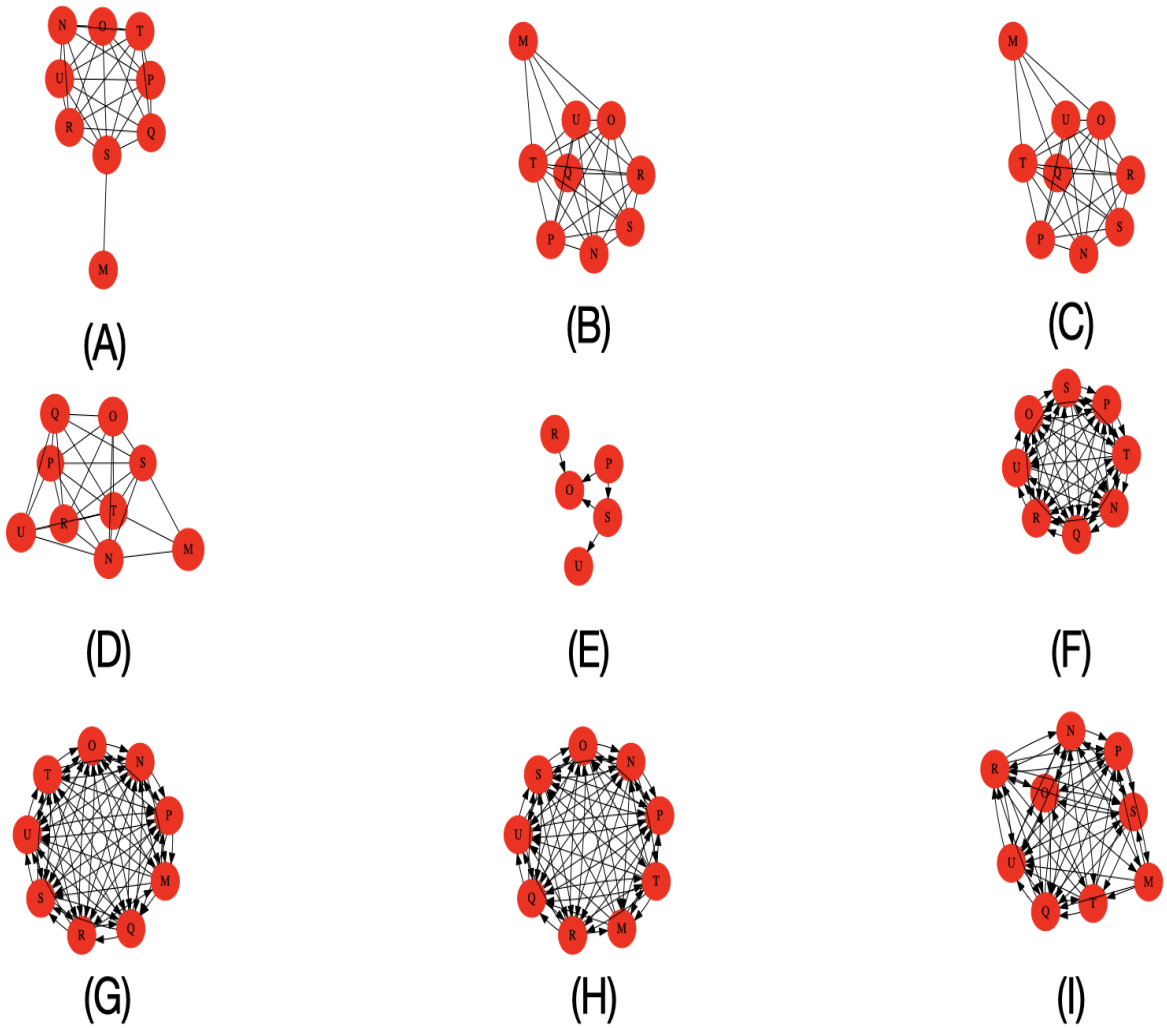

Fig 2. Flight dynamics of homing pigeons. Characterized by (A) Mutual Information(Kernel Estimator) [67] (B) Mutual Information(Discrete & Continuous mixtures) [68] (C) Mutual Information(Kraskov Estimator) [69, 70] (d) Correlation Coefficient (E) Transfer entropy (Kernel estimator) [70] (F) Transfer entropy (Kraskov estimator) [70] (G) Granger Causality (p-value = 0.05) (H) Granger Causality (p-value = 0.001) (I) Granger Causality (p-value = 0.0001). The tables presenting the pair-wise values can be found in S1 Appendix.

There are several potential applications of KLD. For example, it can be used for detecting behavioural change points and modes such as foraging, resting and travelling in animals by constructing a sliding window that moves across a time series while computing the KLD of the probability distributions of contiguous windows. This has implications for example in determining regime shifts, most especially for animals that move in non-homogeneous ways. In addition, it can also be used to identify points of change in landscape for animals who travel long distances over a heterogeneous landscape that affect their behavioural states. KLD has also applications in determining activity-rest patterns in animals.

Another potential application is the use of KLD for monitoring the health of animals

that co-exist in groups. This can be achieved by computing the pairwise KLD between the probability distributions of movement data of all animals in the group while seeking animals with a significant divergence from the remaining members of the group. This approach, when integrated with appropriate machine learning tools such as hierarchical clustering, can be used potentially to classify animals into healthy and non-healthy or automatic classification of animals into species.

3.2 Case study: Kullback-Leibler divergence as measure por predictability of Turkey Vulture annual movement patterns

In this subsection, we demonstrate a potential use of KLD to a biological problem. KLD is indeed widely used as a measure for comparing probability distributions. Fore example, in [21] KLD is used to measure the divergence from the equilibrium behaviour of the blue tuna fish after a telemetry device was attached to it. In this case study, we use KLD to describe the movement patterns of the Turkey Vulture. The Turkey Vulture, is the world's most abundant and widely distributed avian scavenger with a population in excess of five million individuals [40]. We characterize the movement patterns of the Turkey Vulture (*Cathartes aura*) [40, 41] across several years by comparing how the movement patterns at the beginning of the year (January) vary relative to the remaining months of the year across the next three years. The bird by name Leo was chosen since we have movement dataset for several years. We extract up to three years worth of movement data corresponding to 666 data points per month or 28 days and convert the coordinates to distance covered in kilometers. We compute the Kullback-Leibler divergence (symmetric version) across several years and our results (Fig 3) show that the movement strategy of this bird is highly predictable. We observe from Fig 3 that the Kullback-Leibler divergence across the three years is characterized by the presence of several peaks and troughs. The peaks represent the movement back to the breeding sites as well as the breeding season when there is little movement in the temperate regions of America. The troughs represent the migrating period when the birds migrate to tropical regions in search of food. These species of birds start breeding in the temperate regions such as North and South America where they have an abundance of food during the spring and this breeding continues until the onset of fall [40, 42]. In the colder seasons, these birds migrate to tropical regions where it is warmer and there is abundance of rain and food throughout the year. However, at the onset of spring around March, these birds migrate back to the temperate regions of America, where they are guaranteed abundant food and resources. We compare the KL divergence with the mean difference as well as the Earth Mover's Distance (EMD) of the monthly movement data of interest (Fig 3). We also compare how the movement patterns in other months of the year (February to December) vary relative to the remainder of the dataset over three years. The results in Fig 4 show the bird have different movement patterns in the months between June and September relative to other months of the year, which is essentially its breeding season[2].

4 Kolmogorov Complexity

4.1 Overview

Finally, we describe a similarity metric influenced by Kolmogorov complexity (KC), a metric with foundations in the field of algorithmic information theory. The Kolmogorov complexity of an object represents the shortest computer program that produces the

[2] `https://eol.org/pages/45511376/articles`

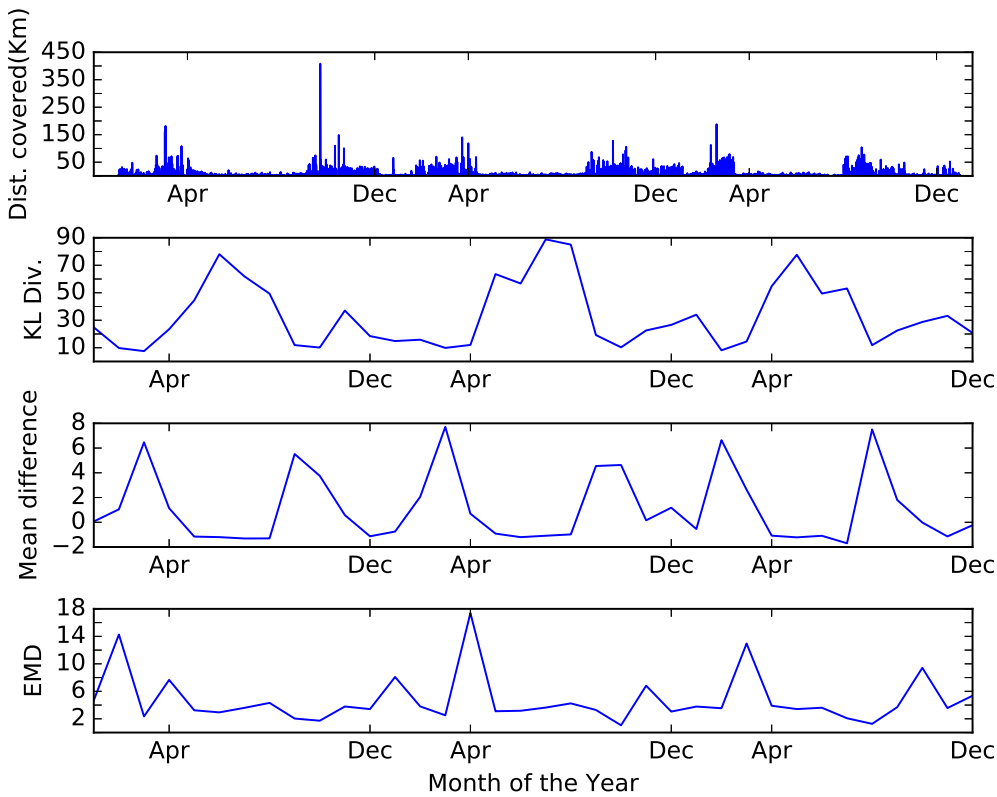

Fig 3. Kullback-Leibler divergence of the monthly movement pattern over a period of 3 years. The peaks represent the annual period of breeding as well as the flight back home after winter and the troughs migratory periods during which the bird travels in search of food. We compare the KL divergence with the mean difference and the Earth Movers Distance (EMD).

object as output [22]. More formally, the KC of a string x with respect to a reference machine U is defined as:

$$\min_{z}\{\{l(z) : U(z) = x, z \in \{0,1\}^*\}. \tag{6}$$

where z is a program that prints string x and then halts and l is the length. The concept of Kolmogorov complexity can be used as an inference tool to find the shortest description of behavioural data. The smaller the KC of a sequence the regular or simple it is and vice-versa. We describe the Normalized Information Distance (NID) [43], a similarity measure inspired by Kolmogorov complexity and defined as:

$$NID(x,y) = \frac{max\ \{K(x|y), K(y|x)\}}{max\ \{K(x), K(y)\}}. \tag{7}$$

Due to the non-computability of the NID, the NID has been re-written [44] as the normalized compression distance (NCD) by simply approximating the Kolmogorov complexity K, using a compressor Z. The NCD between two strings x and y can be defined as:

$$NCD(x,y) = \frac{Z(xy) - min\ \{Z(x), Z(y)\}}{max\ \{Z(x), Z(y)\}}. \tag{8}$$

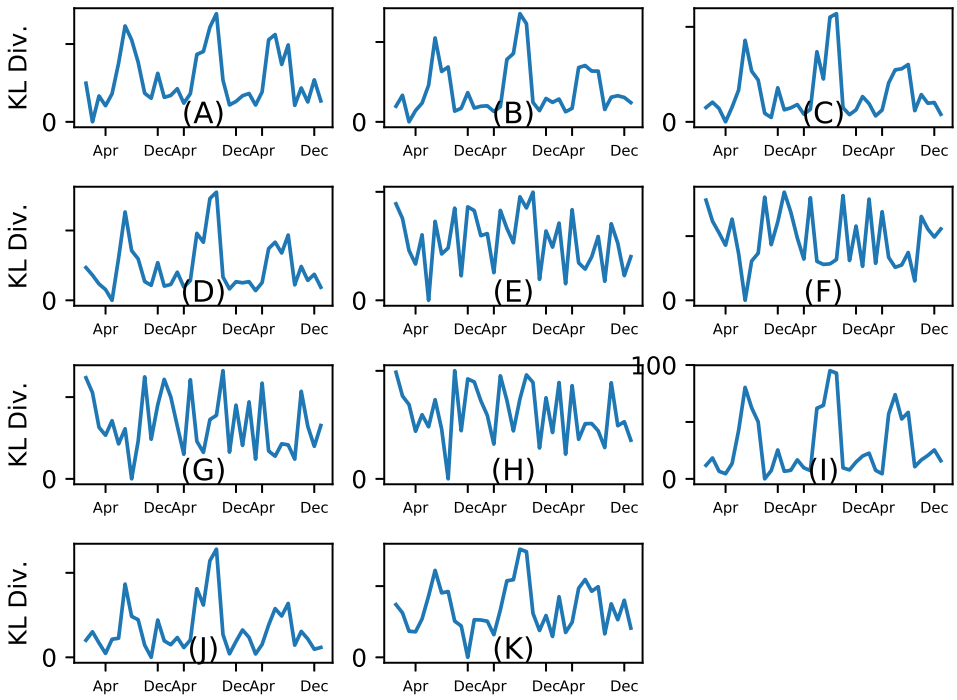

Fig 4. Kullback-Leibler divergence of the monthly movement patterns over a period of 3 years. Where the reference month is (A) February (B) March (C) April (D) May (E) June (F) July (G) August (H) September (I) October (J) November (K) December. From the result, it can be seen that the movement pattern during January, February, March, April, May, October, November and December are the same, while the bird exhibits a different pattern in June, July, August and September. This suggests that the breeding season of this bird is between June and September. The result is consistent with the information about Turkey Vulture in North America as they lay their eggs between May and June, incubate them for between 38 and 41 days and when the eggs hatch. The hatchlings are further brooded for a period between 70 and 80 days resulting in between 108 and 121 days of breeding, which is equivalent to four months of breeding [71].

Here xy is the concatenation of x and y. These strings can be documents, software, genomes or even images. The NCD takes on non-negative values in the range $0 \leq r \leq 1 + \epsilon$ with ϵ defined to take into account imperfections in the compression methods. Please refer to [44] for more details.

The NCD has been used in a variety of disciplines for different purposes, such as anomaly detection [45], gene expression dynamics [7], classification of music [46], detection and classification of computer worms and viruses as well as detecting the origin of new ones [47]. Since the NCD has been shown to work well with sequences and strings, it can be used, for example, in monitoring the behaviour of animals to know when they deviate from a previously or commonly known sequence of states, for example because of climate change [48]. It can also be used to quantify similarity in movement patterns of conspecifics across different habitats.

4.2 Case study: Kolmogorov complexity as tool for classifying animal movement patterns across scales

Animals across different habitats, scales, and species are known to have different movement patterns. However, little or no study has been carried out to find out which groups of animals possess similar movement strategies across different habitats and scales. Recently in [49], the authors discussed a classification of several animals across different species into similar groups using principal component analysis on some movement metrics with hierarchical clustering. Their result suggests that all animals organizes into four distinct groups of movement syndromes namely migratory, central place, nomadic and territorial. In this case study, we analyse the movement patterns of eleven animals (Table 1) across different spatio-temporal scales and habitats. For our analysis, we obtain the datasets of the Galapagos tortoise (*Geochelone nigra*) [50, 51], Springbok (*Antidorcas marsupialis*) [51, 52], African buffalo (*Syncerus caffer*) [51, 53–55], African elephant (*Loxodonta africana*)(original unpublished data contributed by Miriam Tsalyuk and Wayne M. Getz) [51], Black-backed jackal (*Canis mesomelas*) [51, 56], California sea lion (*Zalophus californianus*) (original unpublished data contributed by Dan Costa) [49, 51], Galapagos albatross (*Phoebastria irrorata*) [57], Sheep (*Ovies aries*) and Sheepdog [39, 58], Northern elephant seal (*Mirounga angustirostris*) [51, 59], White-backed vulture (*Gyps africanus*) [51, 60, 61] and Burchell's Zebra (*Equus burchellii*) [62, 63]. All these datasets use the same 1 hour sampling period [49]. First, we evaluate our approach on synthetic data details of which can be found in S1 Appendix. Results, with both short and long movement data shows we can separate distinctly the synthetic animals into groups based on the distribution from which they sample their velocities. Then, we compare the monthly movement patterns of all the 85 animals to find similarities by computing their pairwise NCD with the gzip compressor [64] followed by hierarchical clustering of the resulting distance matrix (see S1 Appendix for details). The only metric used here is distance covered every hour, which we further processed to its binary equivalent (strings of zeroes and ones). We refrained from using the turn-angle here as it is an unreliable metric considering the noisy nature of most sensors. Our results (Fig 5) show three groups emerge: those that live on land (zebra, elephant, springbok, jackal, sheep and buffalo), those that live in water (tortoise, sea lion and elephant seal) and those that fly (Albatross and Turkey Vulture). Amongst the animals that live on land, we notice there appears to be some similarity between the movement patterns of the elephant and zebra while others seem to organize into distinct groups of conspecifics. Therefore, we hypothesize that there might be a correlation between the feeding and movement patterns of animals. We observe a small number of unexpected classifications: for example, Vulture V_1 was classified among the animals that live in water. A possible explanation is related to the fact that the dataset is noisy. To find long-term similarities among animals movement patterns we compare approximately one year movement data of 16 animals (Table 1) across six different species. We selected these datasets due to their temporal length. Results (Fig 6) show that there might be some similarities in the long-term movement patterns of vultures and jackals which are both scavengers. This supports a hypothesis that there might be a correlation in the movement patterns of animals with similar feeding habits.

We compare this approach with the pairwise mutual information of the distance covered with respect to the two instances discussed above (see S1 Appendix). We believe that the insights obtained by means of this technique might be of value for researchers in order to formulate new research hypotheses based on the patterns emerging from this data.

Table 1. Summary of 85 individuals within 11 species used.

Species	No of individuals	Habitat	Feeding strategy
African buffalo (B)	5	Land	Herbivore
African elephant (E)	5(3)	Land	Herbivore
Black-backed jackal (J)	11	Land	Scavengers
Burchell's zebra (Z)	2(2)	Land	Herbivore
California sea lion (sl)	15	Water	Piscivore
Galapagos albatross (A)	8	Land/Air	Piscivore
Galapagos tortoise (T)	7(4)	Land/Water	Piscivore/Herbivore
N. elephant seal (SE)	14	Water	Piscivore
Sheep & Sheep-dog (SD)	5	Land	Herbivore
Springbok (Sp)	9(1)	Land	Herbivore
White-blacked vulture (V)	4(2)	Land/Air	Scavengers

The code in front of the species represents the label codes used for each animals in the hierarchical clustering while the number in the bracket under the no of individuals represents the number of animals with up to one year of observational data.

5 Limitations and Open Challenges

In this paper, we have highlighted the potential use of information theoretic metrics in obtaining insights about animal movement and we have also showed the applications of these metrics to real animal movement data. However, we would like to underline that these methods should be applied with caution given their inherent limitations. First, the issue of missing data remains a challenging problem due to logger failure or inability to regularly obtain position fixes. In addition, it suffices to state here that care must be taken while choosing the appropriate amount of data from which inference can be made. For example, let us consider the analysis of the similarities of the complexities of animals across different taxa and spatio-temporal scales. We identified similarity only in the long-term movement patterns of jackals and vultures. This phenomenon might not have been observed if datasets of shorter length had been used. Indeed the results derived from using the measures presented in this paper should always be evaluated considering the corresponding temporal scale of the dataset.

We issue a caveat on estimating probability density functions (PDF) of continuous movement data. At the moment, most of the methods and tools available are based on the assumption of underlying normal distributions. Considering that continuous animal movement data often follows skewed (e.g., power-law or truncated power-law) distributions [65, 66], researchers employing some of the methods described here, for example the Kullback-Leibler divergence, should exercise appropriate caution while estimating the PDF of these distributions. In the present study, we took special precaution while computing the probability distribution for entropy by binning the data around several mean values of the data in the direction of skewness.

Furthermore, appropriate methods for permutation and randomization must be used while determining a threshold, for example for the calculation of the pairwise mutual information between animals in a group.

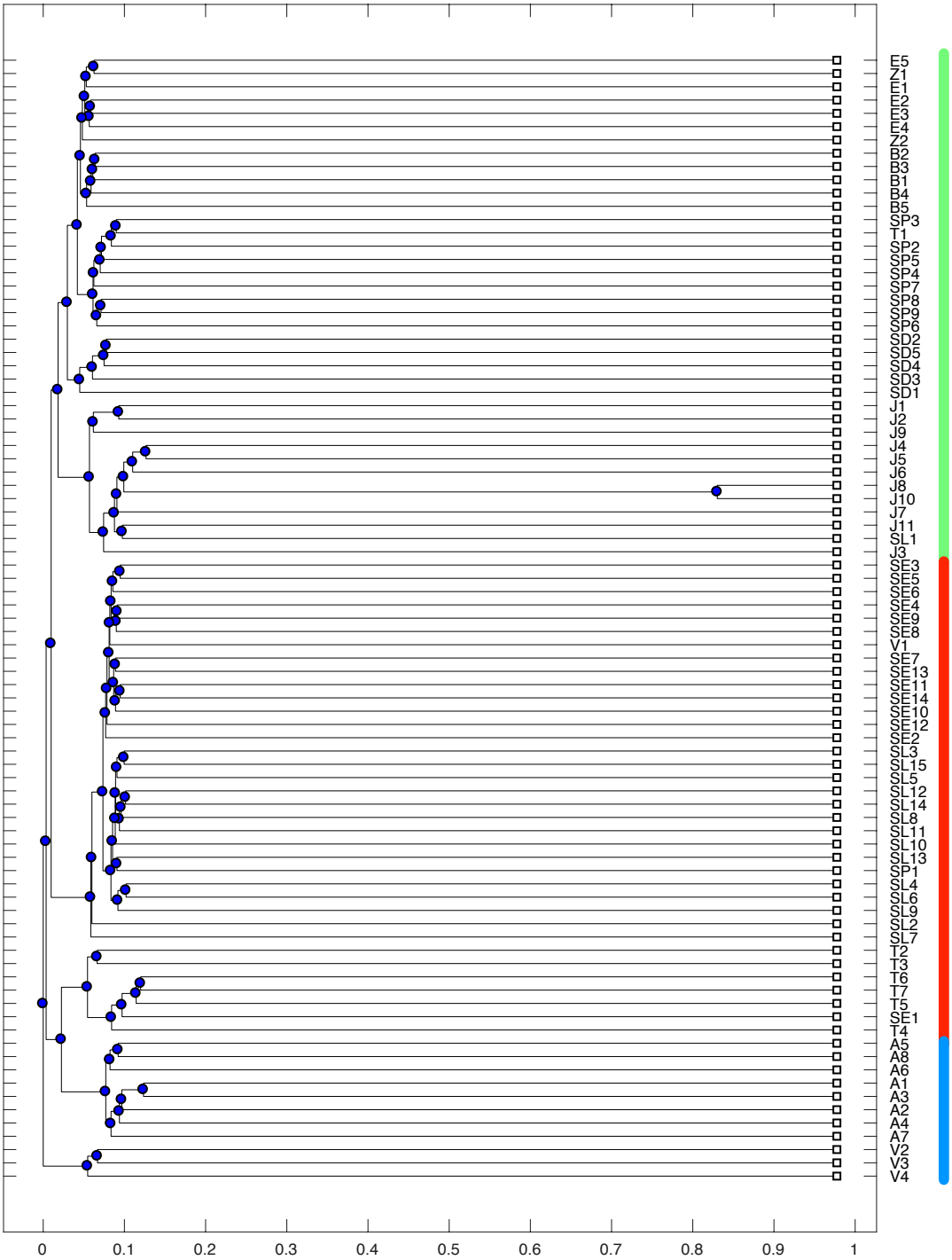

Fig 5. Hierarchical clustering of the pairwise NCD of 85 animals spread across 11 species. All the animals on average organize into three groups of those that live on land (green), those that live in water (red) as well as those who fly (blue).

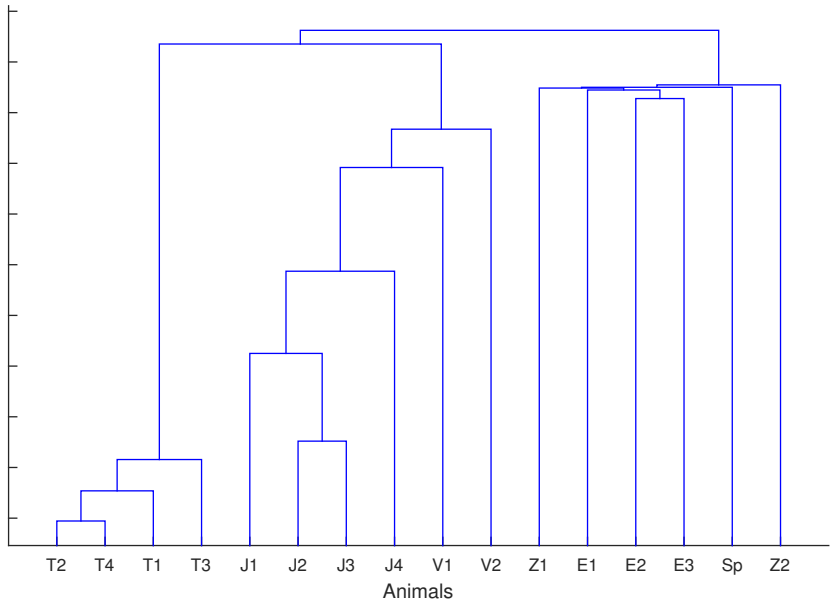

Fig 6. Hierarchical clustering of the pairwise NCD of 16 animals. These animals are spread across 6 species representing the movement patterns over a period of one year. The animals organize into three distinct groups that are correlated with their feeding patterns.

6 Conclusions

In this work, we have demonstrated the use of a class of non-parametric information theoretic tools for studying movement patterns of animals and have showed how they can be applied by means of several animal movement datasets. First, we have demonstrated how Shannon entropy can be used to characterize the movement patterns of sheep with Batten disease where the distance covered every ten minutes was used as the basis for generating symbols to compute the entropy. The result shows that the Batten sheep have a lower entropy than their control counterparts. Then, we have described the use of mutual information for detecting associations in animals using pigeons as an example. Our findings show that this method can be very useful in lieu of the widely used gambit of the group approach as we were able to implicitly detect the leader from the flight data. We have showed how the Kullback-Leibler divergence can be used to characterize the movement patterns of the Turkey Vulture. From our results we were able to to evaluate the predictability of the Turkey vulture over time. Lastly, we have described a metric with foundations in the field of algorithmic information theory known as Kolmogorov complexity (normalized compression distance). We have used this metric to characterize the movement patterns of animals across different taxa and spatio-temporal scales with results suggesting there might be a correlation between the feeding and movement patterns of animals. These methods provide *complementary* insights in the study of animal behaviour. In particular, they can be used to formulate new hypotheses regarding animal movement. In other words, they provide additional information about animal movement that is not apparent using other types of analysis. This class of probabilistic methods is also usually more robust in presence of noise, which is inherent in location data.

As part of our future research agenda, we plan to explore additional movement datasets as they become publicly available and, possibly, other types of behavioural datasets, for example from accelerometers, in order to show further how these information theoretic metrics can be used to obtain novel insights about the behaviour of animals in their natural habitats.

Supporting information

S1 Appendix (PDF)

Acknowledgments

The authors would like to thank Professor Jenny Morton for her insight, valuable discussions, and for allowing them to make use of the sheep dataset in this paper.

Author Contributions

Conceptualization: Kehinde Owoeye, Mirco Musolesi, Stephen Hailes
Data Curation: Kehinde Owoeye, Stephen Hailes
Formal analysis: Kehinde Owoeye
Funding acquisition: Kehinde Owoeye
Investigation: Kehinde Owoeye, Mirco Musolesi, Stephen Hailes
Methodology: Kehinde Owoeye, Mirco Musolesi, Stephen Hailes
Project administration: Mirco Musolesi, Stephen Hailes
Resources: Kehinde Owoeye, Stephen Hailes
Software: Kehinde Owoeye
Supervision: Mirco Musolesi, Stephen Hailes
Validation: Kehinde Owoeye
Visualization: Kehinde Owoeye
Writing - original draft: Kehinde Owoeye
Writing - review & editing: Kehinde Owoeye, Mirco Musolesi, Stephen Hailes

References

1. Gatlin LL, et al. Information theory and the living system. 1972;.
2. Maynard Smith J. The concept of information in biology. Philosophy of science. 2000;67(2):177–194.
3. Shannon CE. A Mathematical Theory of Communication. Bell System Technical Journal. 1948;27(3):379–423. doi:10.1002/j.1538-7305.1948.tb01338.x.
4. Rieke F, Warland D, de Ruyter van R Steveninck, Bialek W. Spikes: Exploring the Neural Code. Cambridge, MA, USA: MIT Press; 1999.
5. Christopher M Bishop. Pattern Recognition and Machine Learning. New York, USA: Springer; 2006.
6. Ueli MM. The role of information theory in cryptography. In: Fourth IMA Conference on Cryptography and Coding; 1993. p. 49–71.

7. Nykter M, Price ND, Aldana M, Ramsey NA, Kauffman SA, Hood LE, et al. Gene expression dynamics in the macrophage exhibit criticality. Proceedings of the National Academy of Sciences. 2008;105(6):1897–1900. doi:10.1073/pnas.0711525105.

8. Macchiavello. Quantum Computation and Quantum Information Theory. Singapore, River Edge, N. J.: World Scientific Publishing; 2000.

9. Lizier JT, Prokopenko M, Zomaya AY. A framework for the local information dynamics of distributed computation in complex systems. In: Guided self-organization: inception. Springer; 2014. p. 115–158.

10. Miller JM, Wang XR, Lizier JT, Prokopenko M, Rossi LF. Measuring information dynamics in swarms. In: Guided self-organization: Inception. Springer; 2014. p. 343–364.

11. King AJ, Wilson AM, Wilshin SD, Lowe J, Haddadi H, Hailes S, et al. Selfish-herd behaviour of sheep under threat. Current Biology. 2012;22(14):R561–R562. doi:http://dx.doi.org/10.1016/j.cub.2012.05.008.

12. Nagy M, Akos Z, Biro D, Vicsek T. Hierarchical group dynamics in pigeon flocks. Nature. 2010;464(7290):890–893. doi:10.1038/nature08891.

13. Mandel JT, Bildstein KL, Bohrer G, Winkler DW. Movement ecology of migration in turkey vultures. Proceedings of the National Academy of Sciences. 2008;105(49):19102–19107. doi:10.1073/pnas.0801789105.

14. Boyland MK, Mlynski DT, James R, Brent LJN, Croft DP. The social network structure of a dynamic group of dairy cows: From individual to group level patterns. Applied Animal Behaviour Science. 2016;174:1–10. doi:http://dx.doi.org/10.1016/j.applanim.2015.11.016.

15. Psorakis I, Roberts SJ, Rezek I, Sheldon BC. Inferring social network structure in ecological systems from spatio-temporal data streams. Journal of the Royal Society Interface. 2012;9(76):3055–3066.

16. Butail S, Ladu F, Spinello D, Porfiri M. Information flow in animal-robot interactions. Entropy. 2014;16(3):1315–1330.

17. Neri D, Ruberto T, Cord-Cruz G, Porfiri M. Information theory and robotics meet to study predator-prey interactions. Chaos: An Interdisciplinary Journal of Nonlinear Science. 2017;27(7):073111.

18. Hu F, Nie LJ, Fu SJ. Information dynamics in the interaction between a prey and a predator fish. Entropy. 2015;17(10):7230–7241.

19. Shannon CE. A mathematical theory of communication. Bell system technical journal. 1948;27(3):379–423.

20. Shannon CE, Weaver W. The mathematical theory of communication. University of Illinois Press; 1949.

21. Kadota M, White EJ, Torisawa S, Komeyama K, Takagi T. Employing Relative Entropy Techniques for Assessing Modifications in Animal Behavior. PLOS ONE. 2011;6(12):1–6. doi:10.1371/journal.pone.0028241.

22. Grunwald P, Vitányi P. Shannon information and Kolmogorov complexity. arXiv preprint cs/0410002. 2004;.

23. Wikelski M, Kays R. Movebank: archive, analysis and sharing of animal movement data. Hosted by the Max Planck Institute for Ornithology; 2017. Available from: `www.movebank.org`.

24. Rènyi A. On Measures of Entropy and Information. In: Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics. University of California Press; 1961. p. 547–561.

25. Tsallis C. Possible generalization of Boltzmann-Gibbs statistics. Journal of Statistical Physics. 1988;52(1):479–487. doi:10.1007/BF01016429.

26. Lord WM, Sun J, Ouellette NT, Bolt EM. Inference of Causal Information Flow in Collective Animal Behavior. IEEE Transactions on Molecular, Biological and Multi-Scale Communications. 2016;2(1):107–116. doi:10.1109/TMBMC.2016.2632099.

27. Perentos N, Martins AQ, Watson TC, Bartsch U, Mitchell NL, Palmer DN, et al. Translational neurophysiology in sheep: measuring sleep and neurological dysfunction in CLN5 Batten disease affected sheep. Brain. 2015;138(4):862–874.

28. Rider JA, Rider DL, Opitz JM, Reynolds JF, Pullarkat RK. Batten disease: past, present, and future. American Journal of Medical Genetics Part A. 1988;31(S5):21–26.

29. Lerner TJ, Boustany RMN, Anderson JW, D'Arigo KL, Schlumpf K, Buckler AJ, et al. Isolation of a novel gene underlying Batten disease, CLN3. Cell. 1995;82(6):949–957.

30. Furmston T, Morton AJ, Hailes S. A Significance Test for Inferring Affiliation Networks from Spatio-Temporal Data. PLOS ONE. 2015;10(7):1–23. doi:10.1371/journal.pone.0132417.

31. Jiang B. Head/Tail Breaks: A New Classification Scheme for Data with a Heavy-Tailed Distribution. The Professional Geographer. 2013;65(3):482–494. doi:10.1080/00330124.2012.700499.

32. Sumpter DJT. Collective Animal Behavior. Princeton University Press; 2010.

33. Krause J, Lusseau D, James R. Animal social networks: an introduction. Behavioral Ecology and Sociobiology. 2009;63(7):967–973. doi:10.1007/s00265-009-0747-0.

34. Franks DW, Ruxton GD, James R. Sampling animal association networks with the gambit of the group. Behavioral Ecology and Sociobiology. 2010;64(3):493–503. doi:10.1007/s00265-009-0865-8.

35. Farine DR. A guide to null models for animal social network analysis. Methods in Ecology and Evolution. 2017;8(10):1309–1320. doi:10.1111/2041-210X.12772.

36. Butail S, Mwaffo V, Porfiri M. Model-free information-theoretic approach to infer leadership in pairs of zebrafish. Physical Review E. 2016;93(4):042411.

37. Santos CD, Neupert S, Lipp HP, Wikeski M, Dechmann DKN. Temporal and Contextual Consistency of Leadership in Homing Pigeon Flocks. PLoS One. 2014;9(7):e102771. doi:10.1371/journal.pone.0102771.

38. Strandburg-Peshkin A, Papageorgiou D, Crofoot MC, Farine DR. Inferring influence and leadership in moving animal groups. Phil Trans R Soc B. 2018;373(1746):20170006.

39. Van Bommel L, Johnson CN. Data from: Where do livestock guardian dogs go? Movement patterns of free-ranging Maremma sheepdogs. Movebank data repository. 2014;doi:10.5441/001/1.pv048q7v.

40. Dodge S, Bohrer G, Bildstein K, Davidson SC, Weinzierl R, Bechard MJ, et al. Environmental drivers of variability in the movement ecology of turkey vultures (Cathartes aura) in North and South America. Philosophical Transactions of the Royal Society of London B: Biological Sciences. 2014;369(1643). doi:10.1098/rstb.2013.0195.

41. Bildstein K, Barber D, Bechard MJ. Data from: Environmental drivers of variability in the movement ecology of turkey vultures (Cathartes aura) in North and South America. Movebank data repository. 2014;doi:10.5441/001/1.46ft1k05.

42. Bildstein KL. Migrating raptors of the world: their ecology & conservation. Cornell University Press; 2006.

43. Vitanyi PMB, Balbach FJ, , Cilibrasi RL, Li M. Information Theory and Statistical Learning:Normalized Information Distance. Boston, MA: Springer US; 2009.

44. Cilibrasi R, Vitanyi P. Clustering by Compression. IEEE Trans Inf Theor. 2005;51(4):1523–1545. doi:10.1109/TIT.2005.844059.

45. Keogh E, Lonardi S, Ratanamahatana CA. Towards Parameter-free Data Mining. In: Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM; 2004. p. 206–215.

46. Cataltepe Z, Yaslan Y, Sonmez A. Music Genre Classification Using MIDI and Audio Features. EURASIP Journal on Advances in Signal Processing. 2007;2007(1):036409. doi:10.1155/2007/36409.

47. Wehner S. Analyzing worms and network traffic using compression. Journal of Computer Security. 2007;15(3):303–320. doi:10.3233/JCS-2007-15301.

48. Hsiung W, Sunstein CR. Climate Change and Animals. University of Pennsylvania Law Review. 2006;155:1695.

49. Abrahms B, Seidel DP, Dougherty E, Hazen EL, Bograd SJ, Wilson AM, et al. Suite of simple metrics reveals common movement syndromes across vertebrate taxa. Movement Ecology. 2017;5(1):12. doi:10.1186/s40462-017-0104-2.

50. Blake S, Yackulic CB, Cabrera F, Tapia W, Gibbs JP, Kummeth F, et al. Vegetation dynamics drive segregation by body size in Galapagos tortoises migrating across altitudinal gradients. Journal of Animal Ecology. 2013;82(2):310–321. doi:10.1111/1365-2656.12020.

51. Abrahms B. Data from: Suite of simple metrics reveals common movement syndromes across vertebrate taxa. Movebank Data Repository. Movebank data repository. 2017;doi:10.5441/001/1.hm5nk220.

52. Andrews JL, Wendy CT, Wayne MG. Home range plus: a space-time characterization of movement over real landscapes. Movement Ecology. 2013;1(1):2. doi:10.1186/2051-3933-1-2.

53. Bar-David S, Bar-David I, Cross PC, Ryan SJ, Knechtel CU, Getz WM. Methods for assessing movement path recursion with application to African buffalo in South Africa. Ecology. 2009;90(9):2467–2479.

54. Cross PC, Bowers JA, Hay CT, Wolhuter J, Buss P, Hofmeyr M, et al. Data from: Nonparameteric kernel methods for constructing home ranges and utilization distributions. Movebank data repository. 2016;doi:10.5441/001/1.j900f88t.

55. Getz WM, Fortmann-Roe S, Cross PC, Lyons AJ, Ryan SJ, Wilmers CC. LoCoH: Nonparameteric kernel methods for constructing home ranges and utilization distributions. PLoS ONE. 2007;2(2). doi:doi:10.1371/journal.pone.0000207.

56. Bellan SE, Cizauskas CA, Miyen J, Ebersohn K, Kusters M, Prager K, et al. Black-backed jackal exposure to rabies virus, canine distemper virus and Bacillus anthracis in Etosha National Park, Namibia. Journal of wildlife diseases. 2012;48(2):371–381. doi:10.7589/0090-3558-48.2.371.

57. Dodge S, Bohrer G, Weinzierl R, Davidson SC, Kays R, Douglas D, et al. The environmental-data automated track annotation (Env-DATA) system: Linking animal tracks with environmental data. Movement Ecology. 2013;1(1):3. doi:10.1186/2051-3933-1-3.

58. Van Bommel L, Johnson CN. Where Do Livestock Guardian Dogs Go? Movement Patterns of Free-Ranging Maremma Sheepdogs. PLoS ONE. 2014;9(10):e111444. doi:10.1371/journal.pone.0111444.

59. Robinson PW, Costa DP, Crocker DE, Gallo-Reynoso JP, Champagne CD, Fowler MA, et al. Foraging Behavior and Success of a Mesopelagic Predator in the Northeast Pacific Ocean: Insights from a Data-Rich Species, the Northern Elephant Seal. PLoS ONE. 2012;7(5):e36728. doi:10.1371/journal.pone.0036728.

60. Spiegel O, Getz WM, Nathan R. Factors Influencing Foraging Search Efficiency: Why Do Scarce Lappet-Faced Vultures Outperform Ubiquitous White-Backed Vultures? The American Naturalist. 2013;181(5):E102–E115. doi:10.1086/655823.

61. Spiegel O, Getz WM, Nathan R. Data from: Factors influencing foraging search efficiency: Why do scarce lappet-faced vultures outperform ubiquitous white-backed vultures? Movebank data repository. 2014;doi:10.5441/001/1.mf903197.

62. Bartlam-Brooks HLA, Beck PSA, Bohrer G, Harris S. In search of greener pastures: Using satellite images to predict the effects of environmental change on zebra migration? Journal of Geophysical Research: Biogeosciences. 2013;118(4):1427–1437. doi:10.1002/jgrg.20096.

63. Bartlam-Brooks HLA, Harris S. Data from: In search of greener pastures: using satellite images to predict the effects of environmental change on zebra migration. Movebank data repository. 2013;doi:10.5441/001/1.f3550b4f.

64. Montero P, Vilar JA. TSclust: An R Package for Time Series Clustering. Journal of Statistical Software. 2014;62(1):1–43.

65. Hirt MR, Jetz W, Rall BC, Brose U. A general scaling law reveals why the largest animals are not the fastest. Nature Ecology & Evolution. 2017;1(8):1116–1122. doi:10.1038/s41559-017-0241-4.

66. Breed GA, Severns PM, Edwards AM. Apparent power-law distributions in animal movements can arise from intraspecific interactions. Journal of the Royal Society Interface. 2015;12(103):1116–1122. doi:10.1098/rsif.2014.0927.

67. Moon YI, Rajagopalan B, Lall U. Estimation of mutual information using kernel density estimators. Physical Review E. 1995;52(3):2318.

68. Gao W, Kannan S, Oh S, Viswanath P. Estimating mutual information for discrete-continuous mixtures. In: Advances in Neural Information Processing Systems; 2017. p. 5986–5997.

69. Kraskov A, Stögbauer H, Grassberger P. Estimating mutual information. Physical review E. 2004;69(6):066138.

70. Lizier JT. JIDT: An information-theoretic toolkit for studying the dynamics of complex systems. Frontiers in Robotics and AI. 2014;1:11.

71. Johnsgard PA. Birds of the Great Plains: Family Cathartidae (American Vultures). 2009;.