

## Research



**Cite this article:** Gelardi V, Godard J, Paleressompoule D, Claidiere N, Barrat A. 2020 Measuring social networks in primates: wearable sensors versus direct observations. *Proc. R. Soc. A* **476**: 20190737. <http://dx.doi.org/10.1098/rspa.2019.0737>

Received: 31 October 2019

Accepted: 28 January 2020

**Subject Areas:**

complexity

**Keywords:**

animal social network, temporal network, wearable proximity sensors

**Author for correspondence:**

Alain Barrat

e-mail: [alain.barrat@cpt.univ-mrs.fr](mailto:alain.barrat@cpt.univ-mrs.fr)

One contribution to a special feature 'A generation of network science'.

Electronic supplementary material is available online at <https://doi.org/10.6084/m9.figshare.c.4915593>.

# Measuring social networks in primates: wearable sensors versus direct observations

Valeria Gelardi<sup>1</sup>, Jeanne Godard<sup>2</sup>, Dany Paleressompoule<sup>2</sup>, Nicolas Claidiere<sup>3</sup> and Alain Barrat<sup>1,4,5</sup>

<sup>1</sup>Aix Marseille Univ, Université de Toulon, CNRS, CPT, Turing Center for Living Systems, Marseille, France

<sup>2</sup>Aix Marseille Univ, CNRS, LPC, FED3C, Marseille, France

<sup>3</sup>Aix-Marseille Univ, Fédération de Recherche 3C, Marseille, France

<sup>4</sup>ISI Foundation, Turin, Italy

<sup>5</sup>Tokyo Tech World Research Hub Initiative (WRHI), Institute of Innovative Research, Tokyo Institute of Technology, Japan

**ID** VG, 0000-0001-6541-2012; NC, 0000-0002-4472-6597;  
AB, 0000-0001-8683-269X

Network analysis represents a valuable and flexible framework to understand the structure of individual interactions at the population level in animal societies. The versatility of network representations is moreover suited to different types of datasets describing these interactions. However, depending on the data collection method, different pictures of the social bonds between individuals could *a priori* emerge. Understanding how the data collection method influences the description of the social structure of a group is thus essential to assess the reliability of social studies based on different types of data. This is however rarely feasible, especially for animal groups, where data collection is often challenging. Here, we address this issue by comparing datasets of interactions between primates collected through two different methods: behavioural observations and wearable proximity sensors. We show that, although many directly observed interactions are not detected by the sensors, the global pictures obtained when aggregating the data to build interaction networks turn out to be remarkably similar. Moreover, sensor data yield a reliable social network over short time scales and can be used for long-term studies, showing their important potential for detailed studies of the evolution of animal social groups.

## 1. Introduction

Interactions between individuals are the foundation of complex social structures in human and other animal societies. Network analysis represents a valuable framework to understand the structure and evolution of these interactions as it encodes a whole hierarchy of patterns, from individual-level interactions to complex population-level social structures [1–8].

With the increasing deployment of digital devices, new ways of collecting data, combined with new network analysis tools, have made possible the development of quantitative measures of these relationships and patterns in modern human societies, leading to the emergence of computational social science over recent decades [9]. For instance, social relationships have been inferred and studied using various data sources ranging from phone calls [10], e-mails [11,12] and online interactions [13] to face-to-face interactions measured by wearable sensors [14–20].

The availability of large volumes of data with high temporal resolution has thus contributed to the rapid expansion of data-driven computational studies of human relationships and human social networks. In contrast, data collection remains more challenging in the field of animal studies, because the data on animal interactions are still largely obtained from direct observations [6,21]. Data resulting from such observations are extremely valuable as they often include detailed information about the nature, duration and location of the interactions between individuals. They thus allow researchers to grasp and investigate complex social patterns in animal groups. Unfortunately, observations are costly in terms of time and, if they are not performed for a long enough time per individual, they have a strong sampling effect [22]. Moreover, observations are almost always biased to some extent because, for example, the visibility of animals is not uniform and some interactions are more easily defined and recognized than others (e.g. [22]).

Recently, work has begun on adapting and implementing a number of technological developments to gather high-resolution behavioural data on non-human animals, leading to the adoption of the term reality mining, which is widely used in computational social sciences for the study of human social behaviour and relations [23], for non-human animal societies [24]. Machine-sensed data relating to the behaviour of animals can indeed now be collected and, most importantly, analysed. We refer to [25] for an overview of the existing and emerging technologies used to collect data on movements, behaviour and interactions within animal groups. In particular, image-based tracking software and machine learning tools can be used to identify and track animals and their trajectories from video data [26–29]. High-resolution GPS can also be used to analyse animals' relative movements: for instance, GPS tracking of wild baboons revealed that a process of shared decision-making governs baboon movements [30]. Different types of data can also be collected jointly (as in the sociometers deployed in human groups [31]); for instance, GPS and audio recordings can be used to investigate the role of vocalization on the cohesion of a group of animals [32].

These recent developments also include proximity logging technologies based on wearable sensors, which are able to provide information either on the distance between the sensor and a fixed receiver [33] or on the distance between two sensors [34–39]. These efforts have enabled the collection of high-resolution datasets in various contexts. Using these techniques presents a number of advantages. First, wearable sensors afford an objective and reliable definition of contact as a proximity event. Second, all individuals equipped with a sensor are monitored together, continuously and potentially for a long time without the need for constant human supervision. This enables, in principle, the collection of large datasets covering long periods of time and, consequently, makes it possible to investigate the evolution and stability of social relationships and social groups on long time scales. On the other hand, wearable sensors do not yield information on the type of behavioural interactions and they do not register contacts with individuals not wearing any sensor, such as very young individuals or members outside the group. The quality of the collected data might also depend on infrastructure constraints and potential technical failures, so that sampling issues must also be considered carefully [20,40].

Thus, data obtained from direct observations and from wearable sensor infrastructures have very different natures and could, in principle, lead to very different descriptions and understanding of the social bonds between individuals and of the resulting social networks [41,42].

It is therefore essential to assess the reliability of the outcomes of social studies to understand the extent to which the way of collecting data influences the final image of the social network, which elements affect this result and how this result is affected. Owing to the difficulty in collecting data using different methods at the same time and in the same population, few studies have been able to address these issues. In human groups, for instance, comparisons between contacts registered by wearable sensors and in diaries have shown both similarities and differences between the data collected by these two methods. In particular, many contacts registered by sensors are not reported in surveys, especially for short contacts, while long contacts are better reported [43–45]. Comparisons between sensors and direct observations or videos have yielded mixed results [46,47]. Among animals, different types of networks built from the same dataset of direct observations have been shown to differ [41,42], while a social network deduced from co-presence in cognitive testing booths has been shown to correlate with that obtained from directly observed interactions [48,49]. However, we are not aware of studies using data collected in the same population with, on the one hand, wearable sensors and, on the other hand, direct observations.

Here, we address this issue by collecting, describing, analysing and comparing two datasets based on dyadic interactions between individuals belonging to a group of Guinea baboons (*Papio papio*). The data span a time window of almost one month between June and July 2019 and were collected through two different methods: (i) behavioural observations by trained human observers and (ii) an infrastructure based on wearable sensors (see <http://www.sociopatterns.org/>).

For these two datasets, we first test the agreement between observations and the sensor data at the level of single events: we systematically check whether an observed interaction was also registered by the sensors. Overall, only a limited fraction of observed interaction events were registered, with strong fluctuations depending on the day of observation and type of behaviour. However, and despite this poor agreement at the level of single events, we show that the time-aggregated networks, and hence the pictures of the group social structure, are remarkably similar.

Finally, we analyse the amount of time that it takes using each data collection method to obtain a robust social network, by comparing the social networks obtained using different time aggregation windows. Strikingly, the social network obtained with just 1 day of sensor data is very similar to that based on aggregation over one whole month of data. Comparatively, the network obtained from observations fluctuates more between short and long aggregation windows because of stronger sampling effects. This shows the potential of wearable sensor infrastructures to detect changes in a social group organization on short time scales and to also monitor its long-term evolution.

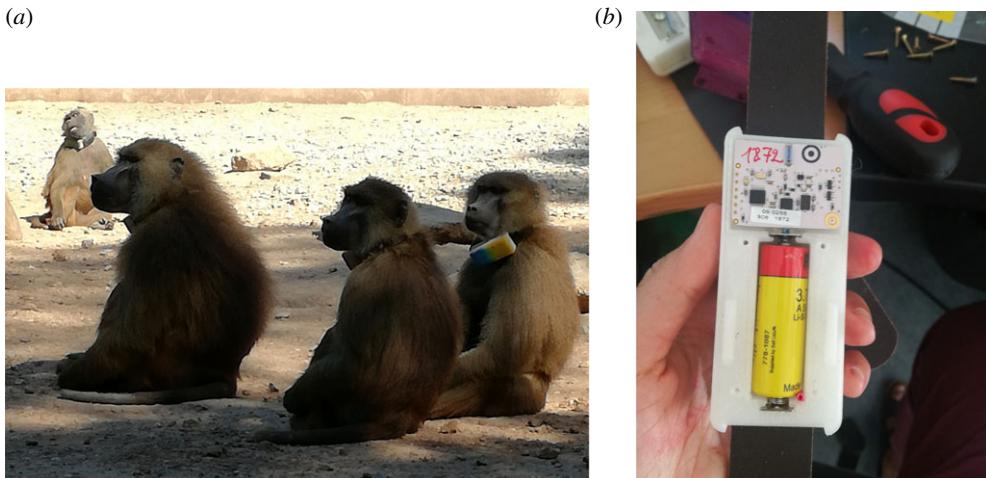
## 2. Methods

### (a) System setting and data collection

Data collection involved a group of captive Guinea baboons living in an enclosure at the CNRS Primate Center in Rousset-sur-Arc (France). The entire group consisted of 19 individuals (seven males and 12 females) aged from 1 to 23 years old.

#### (i) Behavioural observations

The behavioural observations were recorded between 13 June and 10 July 2019 using the focal-sampling method [50]. Observations were carried out for 5 days a week (from Monday to Friday) for a total of 20 days, with two sessions of approximately 2 h a day at different



**Figure 1.** (a) Individuals in the enclosure, wearing collars with the attached boxes containing the tags. (b) Interior of a single box containing the tag (top) and the connected battery (bottom). (Online version in colour.)

times each day, ranging from 08.00 to 17.00. During each session, a trained observer focused on each individual for a period of 5 min and recorded its behaviours. The order in which the different individuals were observed was changed at each session. Fifteen behavioural categories corresponding to social interactions were recorded, namely: 'grooming', 'presenting' (greeting), 'playing with', 'grunting-lipsmacking', 'supplanting', 'threatening', 'submission', 'touching', 'avoiding', 'attacking', 'carrying', 'embracing', 'mounting', 'copulating', 'chasing' (see the electronic supplementary material for details). Another behaviour of interest was 'resting'. Resting is often considered as a behaviour performed in isolation, i.e. it does not correspond to an interaction between individuals. In some cases, however, when two or more individuals were resting together at less than 1 m from each other, this was considered as 'social resting' and also counted as an interaction. 'Resting', 'grooming' and 'playing with' include a duration and are called *state* events. The other types of behaviour do not have a duration assigned to them and are called *point* events. For each observed behaviour, the individuals involved were recorded as well as the start and end time of each state event; for each point event, the time at which it took place was also recorded.

In addition, two categories were included, namely 'invisible' and 'other', to refer to those cases in which the individual was not seen by the observer or the behaviour was not included among those listed above, respectively.

## (ii) Wearable sensor data

A subgroup of 13 baboons, consisting only of juveniles and adults (all individuals at least 6 years old), were fitted with leather collars. Attached to the collars were wearable proximity sensors (radio-frequency identification (RFID) tags) developed by the SocioPatterns collaboration (<http://www.sociopatterns.org/>); these sensors have been used in many studies involving humans [14,16,18,20,40,44,51], and have recently also been used in animals [39]. In our setting, each sensor was secured in a customized box specially designed and produced using a three-dimensional printer to contain the sensor with a long-life battery connected to it. The boxes were positioned on the front side of the individuals (figure 1).

The sensors exchanged low-power radio packets in a peer-to-peer fashion. Owing to the very low power used, the reception by the sensor of an individual A of a radio packet emitted by the sensor of another individual B was a good proxy for the close proximity (approx. less than 1.5 m) of individuals A and B [14]. Moreover, the radio frequency emitted by the RFID tags was absorbed by body water, so the radio packets tended to propagate mostly towards the front

of the individual wearing the device. The packet exchange rate thus depended on the mutual orientation of the individuals and the infrastructure detected mainly face-to-face contacts. The detected spatial proximity relations were relayed from RFID tags to radio receivers (RFID readers) that were installed around the enclosure and connected to a local area network (LAN). A central server received the data, time stamping and storing each event.

Data were finally aggregated with a temporal resolution of 20 s (for more details see [14]): we thus defined two individuals to be in contact during a 20 s time window if their sensors exchanged at least one packet during that interval, and the contact event was considered over when the sensors did not exchange packets over a 20 s interval.

In the following, we will refer to the observed behaviours corresponding to interactions as ‘interactions’ or ‘observed interactions’, and to the contacts collected by the sensors as ‘contacts’ or ‘contact events’.

For the first 10 days of data collection (13–23 June) only two readers were installed around the enclosure, whereas in the successive days a third reader was added to ensure better coverage. Data collection went on even after the observation period was over and is on-going at the time of writing of this paper. We consider here mainly the data collected between 13 June and 10 July 2019, i.e. during the period of the observations; we also use the data collected afterwards and until 27 August to assess stability over longer time scales.

## (b) Data analysis

### (i) Comparison of the two datasets at the level of single events

We compute the fraction of observed interactions that were also detected by the wearable sensor infrastructure as follows. Each observed interaction event involves two individuals  $i$  and  $j$  and is assigned a time  $t$  (for point events) or an interval  $[t_{\text{start}}, t_{\text{stop}}]$  (for state events), with  $t_{\text{start}}$  and  $t_{\text{stop}}$  the times of the beginning and end of the interaction, respectively. Note that the observed interactions are often directed, with an actor individual and a recipient individual. However, since the proximity events registered by the wearable sensors are not directed, we consider undirected versions of the behavioural data (i.e. the direction of the interaction is not taken into account).

An observed interaction event is then considered as *tracked* if at least one RFID packet was exchanged between the sensors of individuals  $i$  and  $j$  within the time window  $[t - \Delta t, t + \Delta t]$  or  $[t_{\text{start}} - \Delta t, t_{\text{end}} + \Delta t]$ , for point and state events, respectively, where  $\Delta t$  is a tolerance interval. This tolerance is introduced to take into account three elements: (i) the potential delay of the observer in reporting the interaction with respect to its actual occurrence; (ii) the 20 s time aggregation of the RFID sensor data; and (iii) possible asynchrony between the time on the observer’s tablet computer on which observed interactions were registered and the time on the computer storing the sensor data.

Obviously, to compute the fraction of tracked interaction events, we only consider observed events involving two individuals wearing sensors.

### (ii) Comparison of the resulting aggregated networks

For each dataset, we can construct on any temporal window an aggregated network in which nodes represent individuals and weighted links give a summary of the recorded contacts or observed interactions during that time window. We first consider the time window as being the whole period during which observations were carried out, and we restrict the observational data to the 13 individuals with collars and sensors. We thus obtain a ‘contact network’ from the wearable sensor data and an ‘interaction network’ from the observed interactions, both covering the period from 27 June to 10 July 2019. Both networks are undirected and weighted.

In the aggregated interaction network, a weighted link between nodes  $i$  and  $j$  is drawn if at least one interaction was observed between  $i$  and  $j$  during the aggregation time window. The weight  $w_{ij}^{(I)}$  of the link between individuals  $i$  and  $j$  is given by the total number of interaction events recorded between  $i$  and  $j$  during this time. Note that we use here the number of interactions and

not their total duration in order to also account for point events. Similarly, in the contact network, a link was drawn between  $i$  and  $j$  if at least one contact was recorded between them by the sensor infrastructure, and the corresponding weight  $w_{ij}^{(C)}$  is given by the number of contacts recorded by the sensors between  $i$  and  $j$ .

We compare the contact and interaction networks using several metrics. We computed the Pearson's and Kendall's  $\tau$  correlation coefficients between the two lists of weights; these are, respectively, the linear correlation and the similarity of the orderings of the weights in the two networks. We also considered two different versions of cosine similarity measures between the network weights. A cosine similarity measure is in general defined between two vectors, and is bounded between  $-1$  and  $+1$ . It takes the value  $+1$  if the vectors are proportional with a positive proportionality constant, a value of  $-1$  if the proportionality constant is negative, and  $0$  if they are perpendicular. For positive weights, as in our case, it is bounded between  $0$  and  $+1$ . We consider first a global cosine similarity (GCS) measure between the two networks as the cosine similarity between the two vectors formed by the list of all link weights in each network (using a weight  $0$  if a link is not present),

$$\text{GCS}_{I,C} = \frac{\sum_{i>j} w_{ij}^{(I)} w_{ij}^{(C)}}{\sqrt{\sum_{i>j} (w_{ij}^{(I)})^2} \sqrt{\sum_{i>j} (w_{ij}^{(C)})^2}}. \quad (2.1)$$

We moreover consider local versions of the cosine similarity: the local cosine similarity (LCS) of a node  $i$  is given by the cosine similarity between the vectors of weights involving  $i$  in each network,

$$\text{LCS}_{I,C}(i) = \frac{\sum_j w_{ij}^{(I)} w_{ij}^{(C)}}{\sqrt{\sum_j (w_{ij}^{(I)})^2} \sqrt{\sum_j (w_{ij}^{(C)})^2}}. \quad (2.2)$$

$\text{LCS}_{I,C}(i)$  is thus equal to  $1$  if  $i$  has been detected as being linked with the same individuals in the two datasets with proportional weights.

It is equal to  $0$  if  $i$  has disjoint sets of neighbours in the two networks. Here, we use the average LCS value over all individuals as a measure of similarity between the two networks.

To get a better grasp of the values obtained, we consider a null model in which the weights are reshuffled among the links for one of the networks. We perform 1000 realizations of this reshuffling and recompute the values of correlations and similarities in each realization, obtaining a null distribution for each measure.

### (iii) Other aggregation time scales: fluctuations and convergence

For each dataset, the aggregation procedure yielding an aggregated network can be performed on any time window. In addition to the whole observation time window, we consider shorter time windows of 1 day, 3 days and one week. We then study the stability and dynamics of each network by computing for each dataset the cosine similarities between the networks aggregated on all pairs of time window (with a given length). For instance, if we consider two different time periods  $t_1$  and  $t_2$ , we denote by  $w_{ij}^{(I,t_1)}$  and  $w_{ij}^{(I,t_2)}$  the weights of the links between individuals  $i$  and  $j$  in the interaction networks aggregated over  $t_1$  and  $t_2$ , respectively, and the LCS of  $i$  in the interaction network between  $t_1$  and  $t_2$  is

$$\text{LCS}_{I,t_1,t_2}(i) = \frac{\sum_j w_{ij}^{(I,t_1)} w_{ij}^{(I,t_2)}}{\sqrt{\sum_j (w_{ij}^{(I,t_1)})^2} \sqrt{\sum_j (w_{ij}^{(I,t_2)})^2}}. \quad (2.3)$$

Local and global cosine similarities can be defined in the same way for each dataset, and can also be defined between networks aggregated on time windows of different lengths. We compute for each dataset similarities between the network aggregated on the whole observation time window and networks aggregated on the first  $n$  days of observations, in order to understand how fast the weight structure of each network converges to its fully aggregated version.

### 3. Results

#### (a) Single interaction and contact events

##### (i) Behavioural data

The total number of behaviours recorded for the entire group of 19 individuals is 5377. From this full dataset of observations we keep just those behaviours involving the 13 individuals that were wearing collars with sensors. Among the 995 observed interactions in this juvenile/adult subgroup, 944 (approx. 95%) are affiliative social behaviours (grooming, resting, presenting, grunting-lipsmacking, touching, mounting, embracing, playing with), which are the most relevant to this study. Moreover, grooming and (social) resting represent more than 98% of the state events and approximately 65% of the total.

Figure 2 represents the distributions of durations for the observed interactions (i.e. with associated duration) both for the subgroup of collared individuals and for the whole group, which includes very young juveniles and babies. Durations cover a broad range of values with a cut-off point at 300 s (5 min) corresponding to the duration of the focal observation (i.e. some interactions lasted more than 300 s but their total duration is unknown). We also note that the distributions of events concerning the whole group or only the collared individuals have similar shapes; however, these have a larger fraction of short interactions when babies are included.

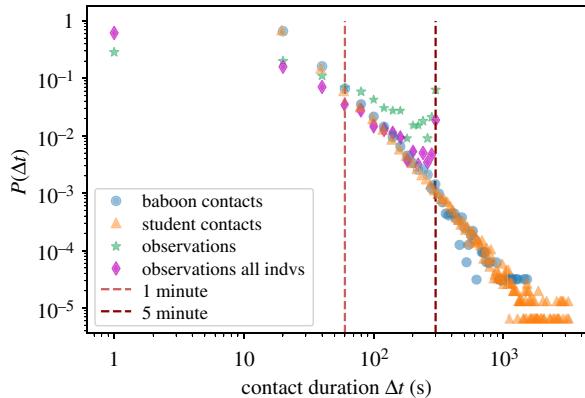
##### (ii) Contacts registered by the sensors

During the period in which observations were carried out, 31 783 contact events were recorded among the 13 individuals. Of these, 4823 (15% of the total) were recorded during the periods of behavioural observations. The number of contacts per day was approximately 1135 on average and ranged from 754 (28 June) to 1768 (13 June) for a total duration of 1 259 500 s (349 h). Contact durations varied across a very broad range: most contacts were short, with an average duration of 39.6 s, and 95% of the contacts lasted less than 2 min, but contacts as long as 1520 s (approx. 25 min) were recorded, and the contact durations form a continuous distribution spanning all values in between (figure 2), as observed in many different contexts for human and animal groups [14,15,18,20,39,44]. In fact, we report on the same graph the statistics of the contact durations measured by wearable sensors between students in a school, reported in [19] and freely accessible: it turns out that the distributions of contact durations of baboons and of humans are indeed very similar.

##### (iii) Comparing interactions and contacts

Figure 2 compares the distributions of durations of the observed interactions and of the contacts registered by the sensor infrastructure. Although they differ, with, in particular, the distribution extending to much larger values for the contacts, they are both broad and span a large range of values. The limited range of the interaction durations is due to the observational protocol, since all durations above 300 s are cut off at that value. In addition, the duration of an interaction is necessarily underestimated if it starts before the beginning of a given 5 min observation time window, or if it terminates after the end of the time window: this biases the resulting distribution of durations in a complex way. Overall, it seems possible that the two distributions would have similar slopes at large durations if this cut-off were not enforced, although a detailed study of the effect of the cut-off is beyond the scope of this work.

To go beyond this statistical comparison, we perform a detailed matching procedure, as described in the Methods, between each single observed interaction (in the behavioural data) and the contact events (obtained from the sensors data). Table 1 gives the results of this matching, for different categories of interactions and different values of the tolerance  $\Delta t$ . The fraction of observed interactions finding a match in the sensor data is quite low, with a slightly better performance when the tolerance is increased. For  $\Delta t = 20$  s, on average only one-third of the



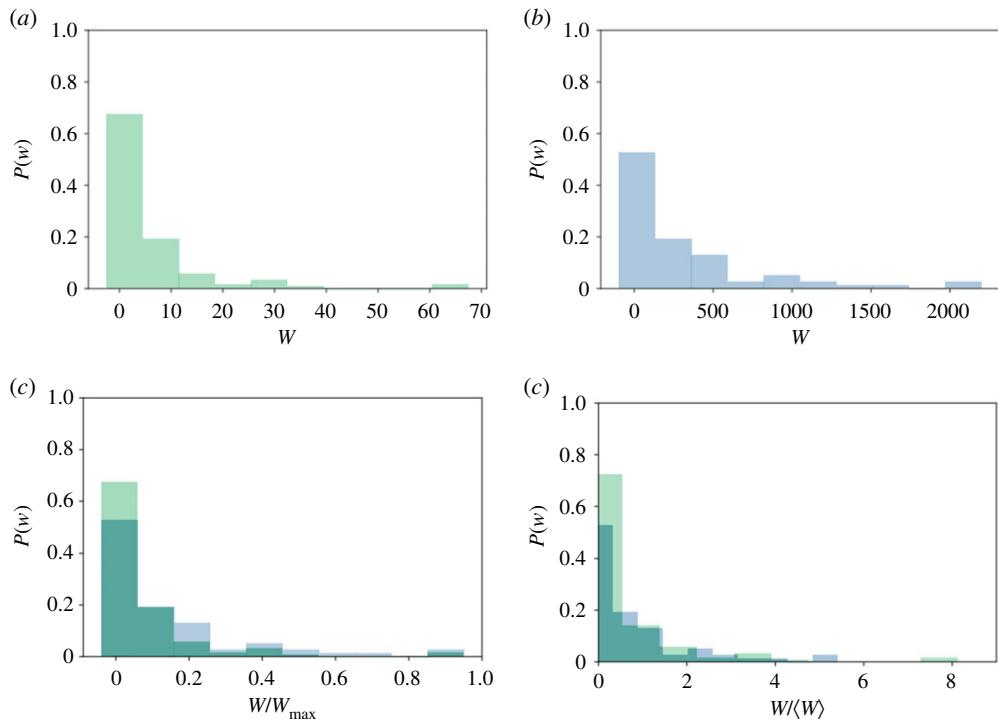
**Figure 2.** Durations of contacts and observed interactions. Distributions of durations for: contacts detected by the wearable sensor infrastructure (blue dots; average = 39.6 s, s.d. = 52.7 s, median = 20.0 s, max = 1520 s); contacts between students in a school in Utah (USA) measured by an infrastructure based on wireless ranging-enabled nodes (WRENs) [19] (orange triangles; average = 39.6 s, s.d. = 72.8 s, median = 19.5 s, max = 3164 s); state events in observed interactions (green stars for interactions involving only the 13 individuals with collars; average = 81.3 s, s.d. = 90.3 s, median = 41.0 s, max = 300 s; magenta diamonds for all interactions; average = 43.8 s, s.d. = 64.6 s, median = 18.0 s, max = 300 s). Note that, according to the observational method used in this study, individuals are observed for 5 min (300 s) at a time. The peak value at 300 s for the observation data is therefore an artefact of the observation method. (Online version in colour.)

**Table 1.** Fractions of observed interactions with a corresponding match in the sensor data. We consider an interaction to have a match in the contact data if the pair of individuals involved in the interaction appear in the sensor data in the same time window  $\pm \Delta t$  (see Methods). We report the overall fraction of matched interactions (first row), the average fraction over the days (second row) with the corresponding standard deviation (third row), and the minimum and the maximum (fourth and fifth row) fractions of tracked interactions over the different days. The values were computed for different delay parameters  $\Delta t$ , and considering either all interactions, only state events (i.e. interactions with duration), only grooming events or only greeting events.

	all interactions			state events			grooming			presenting (greeting)		
$\Delta t$	20 s	40 s	60 s	20 s	40 s	60 s	20 s	40 s	60 s	20 s	40 s	60 s
tot.	0.32	0.36	0.39	0.36	0.4	0.43	0.48	0.53	0.56	0.23	0.27	0.29
avg.	0.32	0.36	0.38	0.37	0.4	0.42	0.49	0.53	0.55	0.2	0.24	0.27
std.	0.12	0.13	0.13	0.12	0.13	0.14	0.19	0.21	0.22	0.2	0.2	0.21
min.	0.1	0.1	0.12	0.11	0.11	0.14	0.17	0.17	0.17	0.0	0.0	0.0
max.	0.54	0.63	0.63	0.61	0.64	0.64	0.83	0.83	0.89	0.67	0.67	0.83

observed interactions appear in the data obtained from the sensor infrastructure. The fraction is lower for very short events such as greetings but notably larger if we consider only grooming events, which are known to be very important socially [52]. We note that in this case the fraction of tracked observed interactions is about 50%, a value very close to that obtained recently in [47] in a comparison between the contacts among human individuals as observed in an annotated video and as registered by wearable sensors.

We finally note that we can consider the reverse procedure that is, considering the contacts registered by the sensors as the ground truth. To this aim, we restrict the sensor contact data to the time windows corresponding to the behavioural observation periods: only 6.63% of these contacts were recorded by the observer as interactions. Note that this small number is not surprising, as



**Figure 3.** Distributions of weights. Probability distributions of interaction (a) and contact (b) network weights. Despite the very different values of the weights, their distributions have similar shapes. In (c) and (d), the same distributions are shown after rescaling the weights of each network by either the maximum weight (c) or the average weight (d). (Online version in colour.)

the observer focuses on one individual at a time, while the sensor infrastructure registers contacts among all collared individuals during the same time window.

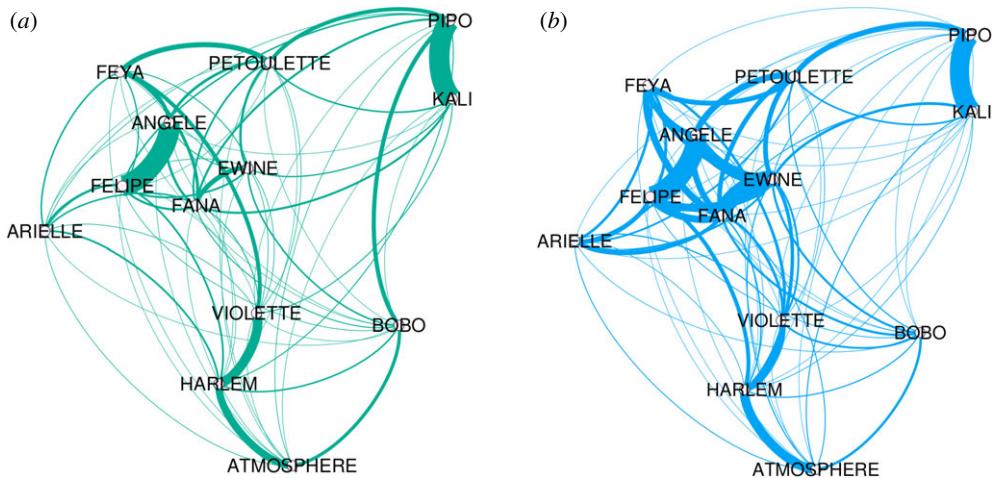
### (b) Comparing interaction and contact networks

Both interaction and contact networks, obtained from the aggregation over the whole period of observation, are very dense, with, respectively, 70 and 78 links (in particular, the contact network is fully connected, i.e. with at least one contact registered within all pairs of individuals).

The two networks have by nature widely different weights, owing to the differences in the methods of measurement. In particular, the number of observed interactions is strongly limited by the amount of time dedicated to the observation of each individual. As a result, the number of observed interactions between a given pair of individuals is at most of a few tens. On the other hand, sensors are active at all times and the weights of the contact network span several orders of magnitude, as is common in such datasets [14]. Despite these differences in the range of weights of the two networks, figure 3 shows that the statistical distributions in fact have very similar shapes.

In addition, figure 4 shows the two weighted networks in which the weights have been rescaled to have comparable widths. This figure highlights some important similarities in the structure of the strong links of both networks: relevant examples include the links Kali–Pipo (female–male) and Angele–Felipe (female–male), as well as the triad Atmosphere–Harlem–Violette (female–male–female) with two similarly strong links Atmosphere–Harlem and Harlem–Violette, and a very weak link Atmosphere–Violette.

To go further, we present in table 2 a systematic comparison of the strongest links in both networks. The table reports the lists of the 10 strongest links in each network. The two strongest links are the same in both networks, and more than half (6/10) of these links appear in both the



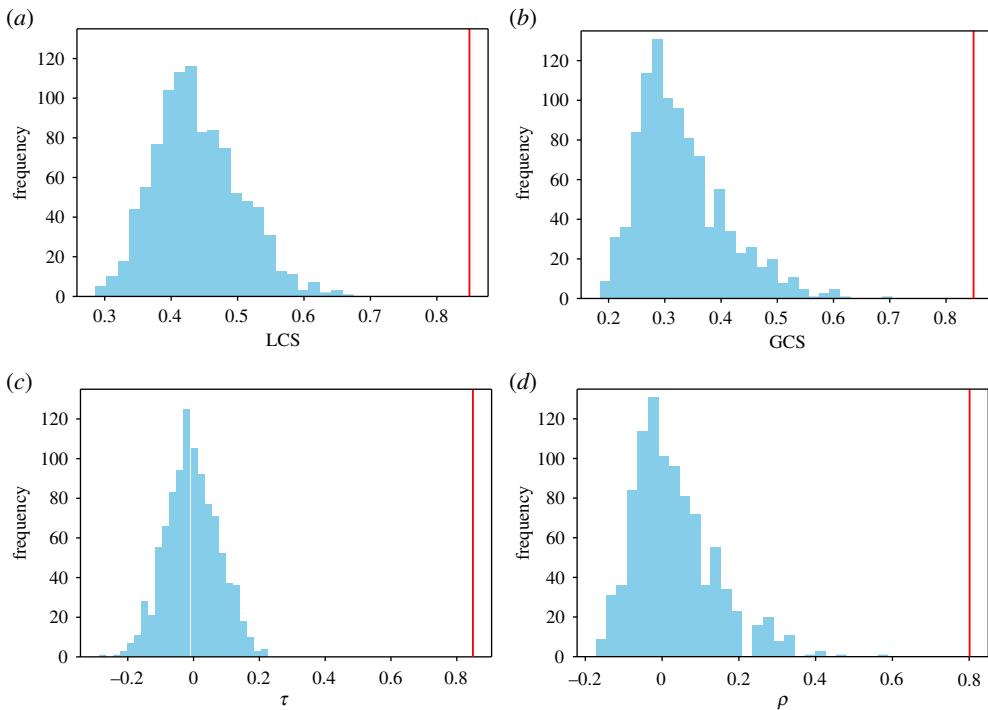
**Figure 4.** Visualization graphs realized with Gephi software ([www.gephi.org](http://www.gephi.org); see also [53]) of the interaction network (a) and the contact network (b), aggregated over the entire observation period. The thickness of the lines is proportional to the weights of the links (scaled in order to have comparable link widths in both networks). The figure shows in a qualitative way the high resemblance between the patterns of strong and weak links of the two networks. (Online version in colour.)

**Table 2.** Top 10 strongest links. For each network, links are ordered based on their weights, from the strongest to the weakest. The links in italics are those present in both top 10 rankings.

rank	interaction network (observations)	contacts network (sensors data)
1	<i>Felipe–Angele</i>	<i>Felipe–Angele</i>
2	<i>Kali–Pipo</i>	<i>Kali–Pipo</i>
3	<i>Violette–Harlem</i>	<i>Fana–Ewine</i>
4	<i>Atmosphere–Harlem</i>	<i>Ewine–Angele</i>
5	<i>Feya–Petoulette</i>	<i>Atmosphere–Harlem</i>
6	<i>Fana–Ewine</i>	<i>Ewine–Felipe</i>
7	<i>Petoulette–Pipo</i>	<i>Violette–Harlem</i>
8	<i>Bobo–Pipo</i>	<i>Fana–Felipe</i>
9	<i>Fana–Felipe</i>	<i>Fana–Angele</i>
10	<i>Ewine–Petoulette</i>	<i>Feya–Angele</i>

interaction and contact networks. The links that are among the top 10 of one network but not of the other are moreover within the top 20 strongest links of the other network, except for one. The exception is the only link between two adult males appearing in table 2, namely the link Bobo–Pipo, ranked eighth for the interaction network but only 55th for the contact network. The interactions between adult males are usually short greetings (presenting), which are not face to face but face to rear (see the description of behaviours in the electronic supplementary material), making this interaction harder for the sensors to detect. We checked that this was indeed the case for the Bobo–Pipo link, with almost only greetings and other point events (92%).

Finally, as described in the Methods, we use four different indicators to give a quantitative estimation of the similarity between the networks. Each indicator is computed for the two empirical networks and for 1000 realizations of the null model described in the Methods, and we compare in figure 5 the empirical value and the distribution of values obtained with the null

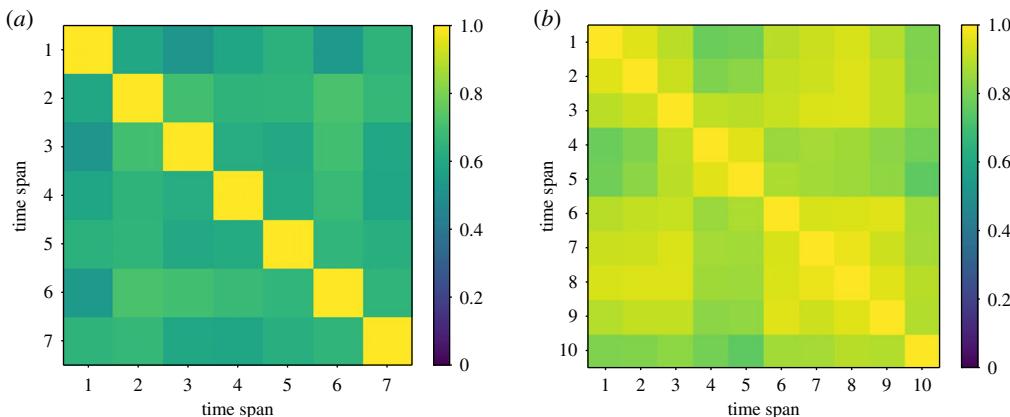


**Figure 5.** Similarity metrics. Quantitative comparison between the contact network and the interaction network through several correlation and similarity measures. In each panel, the empirical value is presented as a vertical red line, together with the distribution of 1000 values (light blue) obtained using a null model in which the weights of the contact network were reshuffled and reassigned at random to the links. (a) Average of the LCS, i.e. average of the cosine similarity between the ego-networks of single nodes (empirical value: 0.91, distribution: average = 0.44, s.d. = 0.06). (b) GCS between the lists of weights of the two networks (empirical value: 0.84, distribution: average = 0.33, s.d. = 0.08). (c) Kendall rank correlation coefficient between the weights of the two networks (empirical value: 0.58, distribution: average = −0.004, s.d. = 0.080). (d) Pearson correlation coefficient between the weights of the two networks (empirical value: 0.80, distribution: average = 0.04, s.d. = 0.11). (Online version in colour.)

model. The GCS and average LCS are extremely high (close to 0.9), as is the Pearson correlation coefficient (0.83), while the Kendall rank correlation coefficient is still large (0.58) but is more impacted by the large number of links with low weights, whose order is expected to be less stable than that of the strong links. In all cases, the empirical values lie far above any value obtained in the null model realizations, as can be clearly seen in the figure (see also the electronic supplementary material for a scatterplot of the weights of the links in the two networks and for the values of the LCS of each individual).

### (c) Other aggregation time scales

As described in the Methods, we also considered other time scales on which to aggregate the data coming from the observations and from the sensor infrastructure. Indeed, the time constraint of observational measures implies that the amount of information concerning each individual per day is relatively small. Building a reliable picture of the social bonds between individuals thus requires many days of observation and an aggregation window of one month is usually advocated [22,48]. In the case of data collected through wearable sensors, on the other hand, a large number of contacts is recorded after only a few hours. However, it is *a priori* unclear whether the structures present on short time scales such as 1 day of data fluctuate from day to day or are



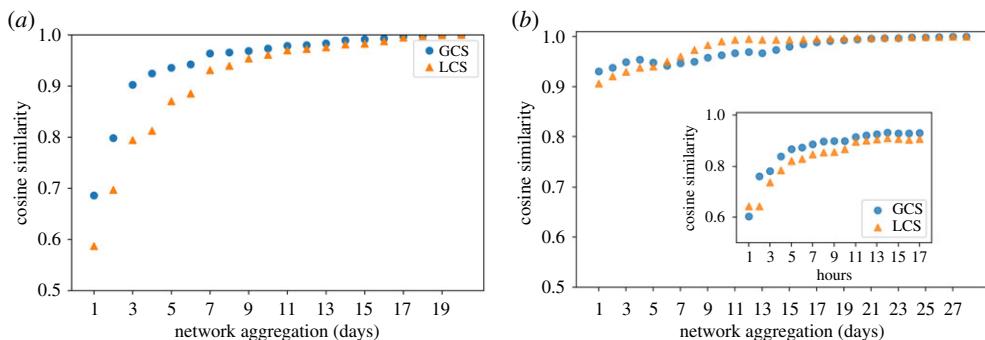
**Figure 6.** Cosine similarities between 3 day networks. Colour-coded matrices of average LCS values between every couple of (a) 3 day interaction networks ( $\min = 0.52$ ;  $\text{mean} = 0.63$ ) and (b) 3 day contact networks ( $\min = 0.75$ ;  $\text{mean} = 0.89$ ). A smaller number of snapshots is obtained for the interaction network because no observations were carried out on Saturdays or Sundays. (Online version in colour.)

stable and already representative of the group social structure and of its strong links. Indeed, the structures present in the monthly aggregated network might result either from a superposition of different daily networks or, on the contrary, from the repetition of the same contact network every day.

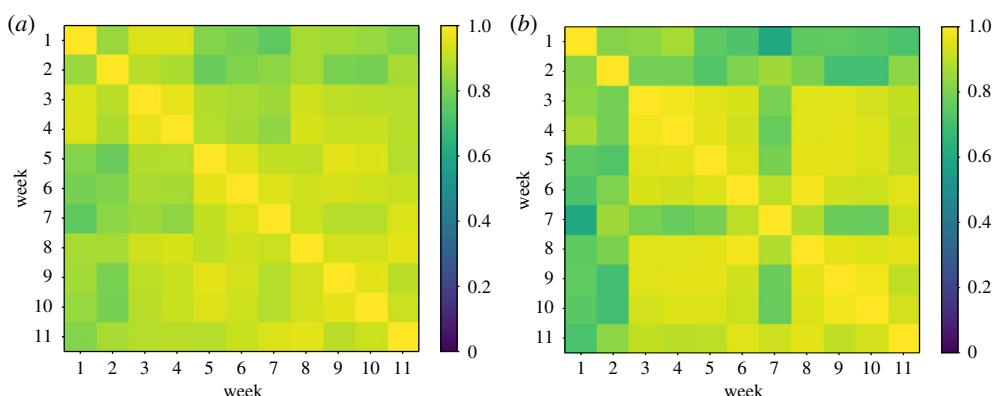
To investigate this issue, we compute the aggregated interaction and contact networks over different time scales (1 day, 3 days, one week), obtaining in each case series of successive snapshots corresponding to the interactions observed or to the contacts measured in successive time windows. For each type of network, we compute the cosine similarities (local and global) between each couple of snapshots to determine how stable the networks are, once aggregated on such time scales. We show in figure 6 the resulting colour-coded matrices for the average LCS: the values obtained for the interaction network are much lower than for the contact network, showing that the former fluctuates much more than the latter. We show in the electronic supplementary material that the interaction networks aggregated at daily scale are even more fluctuating, and that important differences are measured even between weekly aggregated successive networks. On the other hand, daily networks obtained from sensor data are already quite stable.

We then build interaction and contact networks on time windows of increasing lengths, starting from the beginning of the observations (using only the first day of data, then the first 2 days and so on), and comparing them with the aggregated network based on the whole observation time window. Figure 7 shows the resulting GCS and average LCS as a function of the length of the time window considered. The obtained similarities are already close to 1 when only 1 day of sensor data is used ( $\text{GCS} = 0.93$ , average LCS = 0.90), and remain at high values for longer time windows. In fact, the inset shows that similarity values rise above 0.8 as soon as about 10–12 h of data are collected. Comparatively, the similarities to the final aggregated network increase much more slowly for the interaction data than for the contact data. For instance, the 1 day interaction network has much less similarity to the monthly one ( $\text{GCS} = 0.68$ , average LCS = 0.58). However, they reach very high values after only 9–10 days of observation.

Finally, to illustrate the possibility of exploring long time scales using the sensor infrastructure, we consider the data collected after the observation period was over. Figure 8 shows the GCS and average LCS between networks aggregated on weekly time scales, from 13 June to 27 August 2019. The picture emerging from figure 8 is one of a very stable network, with average similarity values of 0.89 and 0.87 for the average LCS and for the GCS, respectively. Note that the minima of the similarities between weekly networks (0.75 and 0.61) are observed for the week 18–25 July,



**Figure 7.** Global and average local similarities between the networks aggregated on the whole period of observation and networks aggregated on shorter time windows. (a) Interaction networks and (b) contact networks. (Online version in colour.)



**Figure 8.** Stability of weekly aggregated contact networks on long time scales. Colour-coded matrices of the average LCS values (a) and GCS values (b) for all pairs of weekly contact networks from 13 June to 27 August. (Online version in colour.)

in which the infrastructure actually failed for a couple of days, resulting in data loss for 20 and 21 July.

## 4. Discussion

In this paper, we have analysed and compared datasets describing social relationships in a group of non-human primates, collected through two different methods: behavioural observations and wearable proximity sensors. Sensor and behavioural observation methods have different advantages and limitations that influence their ability to detect contacts and interactions between individuals. On the one hand, observational methods provide high-quality data, through which we can distinguish different behaviours and describe in depth social relationships between individuals. However, the data have several sampling issues: only a limited amount of time can be spent observing each individual, and data may be completely absent on certain days for logistical reasons (Saturdays and Sundays in our case). Biases related to the observation technique can also occur and are difficult to estimate. The cut-off on the duration of each individual observation leads moreover to an underestimation of the duration of long interactions, which can be particularly important when determining the social structure of a group. Finally, the total duration of the observation period is usually limited to a few weeks and the group cannot be monitored continuously on very long time scales, for clear practical reasons.

On the other hand, sensors provide a large amount of data in a continuous manner, with a high temporal resolution and potentially on very long time scales. However, the advantage of having an objective definition of a contact as an exchange of radio packets between sensors is at the same time a limitation since no information on the behaviour of the individuals in contact is available. In particular, it is not possible to distinguish affiliative behaviours from antagonistic ones, and it is therefore not possible to determine the dominance hierarchy using only wearable proximity sensors. Moreover, two types of sampling biases need to be mentioned. First, the directionality of the sensors limits the detection to approximately face-to-face interactions, while some social interactions between primates (in particular, greetings or social resting) occur when individuals have different mutual orientations. Second, individuals not wearing the sensors are by definition absent from the data. In the case of human populations, it has been shown that a uniform population sampling does not alter the statistical properties of the contact network between individuals [20]. However, the absence of data concerning a specific subgroup of individuals with behavioural patterns different from the rest of the group (such as all the individuals less than 6 years old, in our case) is a clear limitation. Another difficulty could also come from data losses whenever a part of the infrastructure fails. We show, however, in the electronic supplementary material by simulating the failure of a reader that the structure of the social network deduced from the sensor data remains very stable even when the amount of data lost is important.

We have performed a comparison of the two datasets at various levels of detail. At the most detailed level, we have checked for each observed interaction whether a contact was registered by the sensors at the same time. This matching turns out to be very limited: about one-third of the observed interactions were also recorded by the sensors, but this amount fluctuates and depends strongly on the type of interaction. In particular, for interactions that tend to last, such as grooming, the percentage of tracked events rises to approximately 50% and is above approximately 80% on some days of observations. This is particularly important as grooming behaviour is for primates one of the core social interactions, allowing the social structure of the group to be defined [52]. For short and elusive interactions instead, such as greeting, this percentage is only approximately 20%, which can be explained by the fact that greetings among baboons are most often not face-to-face interactions. Notably, our results are in line with [47], where the correspondence was examined between data collected by wearable sensors and a video of the same interactions, yielding a sensitivity of 50% (about half of the interactions annotated on the video were present in the wearable sensor data).

Although this limited correspondence between the two methods of measuring interactions could be seen as a negative result, it is striking that, when considering the global social structure extracted from the data, the results of the two methods are in fact strikingly similar. First, at a statistical level, the distributions of event durations are broad in both cases, with most events having a short duration, and a continuously decreasing distribution with no cut-off except the one imposed by the procedure. The distributions of weights (number of events between two individuals) are also very similar. Most importantly, the networks aggregated over the whole observation period turn out to be extremely similar as measured by several metrics: the weights of a link joining two individuals are highly correlated in the two networks, the top-ranked links and the strong structures are preserved. Overall, the pictures of the social network provided by the two measurement systems are thus extremely similar, despite the discrepancies observed at the very detailed level. We note that this result is at odds with the analysis of [41], in which an interaction network and a proximity network, built from the same set of direct observations of baboons, were shown to differ. However, our infrastructure detects very close proximity, which would be allowed only between animals sharing a certain level of trust, while the proximity criterion used by Castles *et al.* [41] was of 5 or 10 m, thus not corresponding to a ‘contact’ between individuals.

Moreover, we have built aggregated networks at shorter time scales and investigated how quickly a reliable network could be obtained in each case. The observation network fluctuates strongly from day to day. It yields a very similar picture with respect to the whole observation period after about 10 days of observation. This is in agreement with recent results showing that

limited amounts of observational data were enough to obtain a clear picture of the social structure of the group [22]. Comparatively, the structure of the aggregated network obtained with the sensors is already very stable at short time scales, and we obtain a reliable network structure very similar to the one aggregated over one month even with only 1 day of data. This implies that the sensor data would potentially allow us to pinpoint a change in the structure of the social network of the group on much shorter time scales than with observational data. Moreover, we have actually deployed the sensor infrastructure on much longer time scales with no interruption (currently up to six months), while continuous observations cannot be carried out realistically on such time scales.

## 5. Conclusion

Collecting and analysing data coming from digital devices to evaluate social patterns has become quite common in human social studies. Recently, new infrastructures and protocols based on sensors have also become more easily available for the study of animal groups as an alternative to traditional data collection methods such as behavioural observations. In particular, we have shown the potential of the sensor infrastructure used here, which detects close proximity between individuals, to enable an automatic, less costly, long-term and reliable data collection that yields a picture of the social interactions very similar to the one obtained from direct observations, despite not registering all observed interactions and not distinguishing between different types of behaviours. Such techniques could thus facilitate animal social network analysis and most importantly make accessible both short and long time scales for the investigation of the dynamical evolution of animal social networks [54].

**Ethics.** The baboons lived in an outdoor enclosure ( $700\text{ m}^2$ ) connected to an indoor area that provided shelter when necessary. Water was provided ad libitum within the enclosure, and they received their normal ratio of food (fruits, vegetables and monkey chow) every day at 17.00. The baboons were all born within the primate centre. This research was carried out in accordance with French standards and received approval from the national French ethics committee, the ‘Comité d’Ethique CE-14 pour l’Expérimentation Animale’ (approval no. APAFIS#4816-2015091110584769).

**Data accessibility.** The data that support the findings of this study are openly available in the Open Science Foundation repository at doi:10.17605/OSF.IO/UFS3Y.

**Authors’ contributions.** A.B. and N.C. designed the study. A.B., N.C., D.P. and V.G. adapted the sensors for the baboons. V.G. and J.G. carried out the data collection. V.G. performed the data analysis and wrote the first draft of the manuscript. A.B., N.C. and V.G. wrote the final version of the manuscript. All authors gave final approval for publication and agree to be held accountable for the work performed therein.

**Funding.** This work was supported by Agence Nationale de la Recherche grants to N.C. (no. ANR ASCE ANR-13-PDOC-0004) and to A.B. (DATAREDUX ANR-19-CE46-0008-01). The funders had no role in this study.

**Acknowledgements.** The authors thank the staff at the Roussel-sur-Arc Primate Center (CNRS-UPS846, France) for making the study possible and Julie Gullstrand and Joël Fagot for helping in data collection.

## References

1. Hinde RA. 1976 Interactions, relationships and social structure. *Man* **11**, 1–17. (doi:10.2307/2800384)
2. Croft D, James R, Krause J. 2008 *Exploring animal social networks*. Princeton, NJ: Princeton University Press.
3. Wey T, Blumstein DT, Shen W, Jordán F. 2008 Social network analysis of animal behaviour: a promising tool for the study of sociality. *Anim. Behav.* **75**, 333–344. (doi:10.1016/j.anbehav.2007.06.020)
4. Hobson EA, Avery ML, Wright TF. 2013 An analytical framework for quantifying and testing patterns of temporal dynamics in social networks. *Anim. Behav.* **85**, 83–96. (doi:10.1016/j.anbehav.2012.10.010)

5. Kurvers RHJM, Krause J, Croft DP, Wilson ADM, Wolf M. 2014 The evolutionary and ecological consequences of animal social networks: emerging issues. *Trends Ecol. Evol.* **29**, 326–335. (doi:10.1016/j.tree.2014.04.002)
6. Krause J, James R, Franks D, Croft D. 2015 *Animal social networks*. Oxford, UK: Oxford University Press.
7. Whitehead H. 2008 *Analyzing animal societies: quantitative methods for vertebrate social analysis*. Chicago, IL: University of Chicago Press.
8. Webber QM, Wal EV. 2019 Trends and perspectives on the use of animal social network analysis in behavioural ecology: a bibliometric approach. *Anim. Behav.* **149**, 77–87. (doi:10.1016/j.anbehav.2019.01.010)
9. Wal EV *et al.* 2009 Life in the network: the coming age of computational social science. *Science* **323**, 721–723. (doi:10.1126/science.1167742)
10. Onnela JP, Saramäki J, Hyvonen J, Szabó G, Argollo de Menezes M, Kaski K, Barabási AL, Kertész J. 2007 Analysis of a large-scale weighted network of one-to-one human communication. *New J. Phys.* **9**, 179–179. (doi:10.1088/1367-2630/9/6/179)
11. Eckmann JP, Moses E, Sergi D. 2004 Entropy of dialogues creates coherent structures in e-mail traffic. *Proc. Natl Acad. Sci. USA* **101**, 14 333–14 337. (doi:10.1073/pnas.0405728101)
12. Kossinets G, Watts DJ. 2006 Empirical analysis of an evolving social network. *Science* **311**, 88–90. (doi:10.1126/science.1116869)
13. Golder SA, Wilkinson DM, Huberman BA. 2007 Rhythms of social interaction: messaging within a massive online network. In *Communities and technologies 2007* (eds C Steinfield, BT Pentland, M Ackerman, N Contractor), pp. 41–66. Berlin, Germany: Springer.
14. Cattuto C, Van den Broeck W, Barrat A, Colizza V, Pinton JF, Vespignani A. 2010 Dynamics of person-to-person interactions from distributed RFID sensor networks. *PLoS ONE* **5**, e11596. (doi:10.1371/journal.pone.0011596)
15. Salathé M, Kazandjieva M, Lee JW, Levis P, Feldman MW, Jones JH. 2010 A high-resolution human contact network for infectious disease transmission. *Proc. Natl Acad. Sci. USA* **107**, 22 020–22 025. (doi:10.1073/pnas.1009094108)
16. Isella L, Stehlé J, Barrat A, Cattuto C, Pinton JF, Van den Broeck W. 2011 What's in a crowd? Analysis of face-to-face behavioural networks. *J. Theor. Biol.* **271**, 166–180. (doi:10.1016/j.jtbi.2010.11.033)
17. Stopczynski A, Sekara V, Sapiezynski P, Cuttone A, Madsen MM, Larsen JE, Lehmann S. 2014 Measuring large-scale social networks with high resolution. *PLoS ONE* **9**, e95978. (doi:10.1371/journal.pone.0095978)
18. Fournet J, Barrat A. 2014 Contact patterns among high school students. *PLoS ONE* **9**, e107878. (doi:10.1371/journal.pone.0107878)
19. Toth DJ, Leecaster M, Pettey WB, Gundlapalli AV, Gao H, Rainey JJ, Uzicanin A, Samore MH. 2015 The role of heterogeneity in contact timing and duration in network models of influenza spread in schools. *J. R. Soc. Interface* **12**, 20150279. (doi:10.1098/rsif.2015.0279)
20. Génoin M, Vestergaard C, Cattuto C, Barrat A. 2015 Compensating for population sampling in simulations of epidemic spread on temporal contact networks. *Nat. Commun.* **8**, 8860. (doi:10.1038/ncomms9860)
21. Farine DR, Whitehead H. 2015 Constructing, conducting and interpreting animal social network analysis. *J. Anim. Ecol.* **84**, 1144–1163. (doi:10.1111/1365-2656.12418)
22. Davis GH, Crofoot MC, Farine DR. 2018 Estimating the robustness and uncertainty of animal social networks using different observational methods. *Anim. Behav.* **141**, 29–44. (doi:10.1016/j.anbehav.2018.04.012)
23. Eagle N, Pentland A. 2006 Reality mining: sensing complex social systems. *Personal Ubiquitous Comput.* **10**, 255–268. (doi:10.1007/s00779-005-0046-3)
24. Krause J, Krause S, Arlinghaus R, Psorakis I, Roberts S, Rutz C. 2013 Reality mining of animal social systems. *Trends Ecol. Evol.* **28**, 541–551. (doi:10.1016/j.tree.2013.06.002)
25. Hughey LF, Hein AM, Strandburg-Peshkin A, Jensen FH. 2018 Challenges and solutions for studying collective animal behaviour in the wild. *Phil. Trans. R. Soc. B* **373**, 20170005. (doi:10.1098/rstb.2017.0005)
26. Crall J, Gravish N, Mountcastle A, Combes S. 2015 BEEtag: a low-cost, image-based tracking system for the study of animal behaviour and locomotion. *PLoS ONE* **10**, e0136487. (doi:10.1371/journal.pone.0136487)

27. Romero-Ferrero F, Bergomi MG, Hinz RC, Heras FJH, de Polavieja GG. 2019 idtracker.ai: tracking all individuals in small or large collectives of unmarked animals. *Nat. Methods* **16**, 179–182. (doi:10.1038/s41592-018-0295-5)
28. Mathis A, Mamidanna P, Cury KM, Abe T, Murthy VN, Mathis MW, Bethge M. 2018 DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. *Nat. Neurosci.* **21**, 1281–1289. (doi:10.1038/s41593-018-0209-y)
29. Graving JM, Chae D, Naik H, Li L, Koger B, Costelloe BR, Couzin ID. 2019 DeepPoseKit, a software toolkit for fast and robust animal pose estimation using deep learning. *eLife* **8**, e47994. (doi:10.7554/eLife.47994)
30. Strandburg-Peshkin A, Farine DR, Couzin ID, Crofoot MC. 2015 Shared decision-making drives collective movement in wild baboons. *Science* **348**, 1358–1361. (doi:10.1126/science.aaa5099)
31. Choudhury T, Pentland A. 2003 Sensing and modeling human networks using the sociometer. In *Proc. of the 7th IEEE Int. Symp. on Wearable Computers ISWC '03, White Plains, NY, 21–23 October 2003*. Washington, DC: IEEE Computer Society.
32. O'Bryan LR, Abaid N, Nakayama S, Dey T, King AJ, Cowlishaw G, Rubenstein DI, Garnier S. 2019 Contact calls facilitate group contraction in free-ranging goats (*Capra aegagrus hircus*). *Front. Ecol. Evol.* **7**, 73. (doi:10.3389/fevo.2019.00073)
33. Ryder TB, Horton BM, Morales JDD, Moore IT. 2012 Proximity data-loggers increase the quantity and quality of social network data. *Biol. Lett.* **8**, 917–920. (doi:10.1098/rsbl.2012.0536)
34. Rutz C, Burns ZT, James R, Ismar SM, Burt J, Otis B, Bowen J, St Clair JJ. 2012 Automated mapping of social networks in wild birds. *Curr. Biol.* **22**, R669–R671. (doi:10.1016/j.cub.2012.06.037)
35. St Clair JJ, Burns ZT, Bettaney EM, Morrissey MB, Otis B, Ryder TB, Fleischer RC, James R, Rutz C. 2015 Experimental resource pulses influence social-network dynamics and the potential for information flow in tool-using crows. *Nat. Commun.* **6**, 7197. (doi:10.1038/ncomms8197)
36. Rutz C, Morrissey MB, Burns ZT, Burt J, Otis B, St Clair JJ, James R. 2015 Calibrating animal-borne proximity loggers. *Methods Ecol. Evol.* **6**, 656–667. (doi:10.1111/2041-210X.12370)
37. Bettaney E, James R, St Clair J, Rutz C. 2015 Processing and visualising association data from animal-borne proximity loggers. *Anim. Biotelem.* **3**, 27. (doi:10.1186/s40317-015-0065-4)
38. Milwid RM, O'Sullivan TL, Poljak Z, Laskowski M, Greer AL. 2019 Validation of modified radio-frequency identification tag firmware, using an equine population case study. *PLoS ONE* **14**, e0210148. (doi:10.1371/journal.pone.0210148)
39. Wilson-Agarwal JK, Ozella L, Tizzoni M, Cattuto C, Swan GJ, Moundai T, Silk MJ, Zingeser JA, McDonald RA. 2019 High-resolution contact networks of free-ranging domestic dogs *Canis familiaris* and implications for transmission of infection. *PLoS Negl. Trop. Dis.* **13**, e0007565. (doi:10.1371/journal.pntd.0007565)
40. Vestergaard CL, Valdano E, Génois M, Poletto C, Colizza V, Barrat A. 2016 Impact of spatially constrained sampling of temporal contact networks on the evaluation of the epidemic risk. *Eur. J. Appl. Math.* **27**, 941–957. (doi:10.1017/S0956792516000309)
41. Castles M, Heinsohn R, Marshall HH, Lee AE, Cowlishaw G, Carter AJ. 2014 Social networks created with different techniques are not comparable. *Anim. Behav.* **96**, 59–67. (doi:10.1016/j.anbehav.2014.07.023)
42. Farine DR. 2015 Proximity as a proxy for interactions: issues of scale in social network analysis. *Anim. Behav.* **104**, e1–e5. (doi:10.1016/j.anbehav.2014.11.019)
43. Smieszek T, Barclay VC, Seenii I, Rainey JJ, Gao H, Uzicanin A, Salathé M. 2014 How should social mixing be measured: comparing web-based survey and sensor-based methods. *BMC Infect. Dis.* **14**, 136. (doi:10.1186/1471-2334-14-136)
44. Mastrandrea R, Fournet J, Barrat A. 2015 Contact patterns in a high school: a comparison between data collected using wearable sensors, contact diaries and friendship surveys. *PLoS ONE* **10**, e0136497. (doi:10.1371/journal.pone.0136497)
45. Smieszek T, Castell S, Barrat A, Cattuto C, White PJ, Krause G. 2016 Contact diaries versus wearable proximity sensors in measuring contact patterns at a conference: method comparison and participants' attitudes. *BMC Infect. Dis.* **16**, 341. (doi:10.1186/s12879-016-1676-y)
46. Lucet JC *et al.* 2012 Electronic sensors for assessing interactions between healthcare workers and patients under airborne precautions. *PLoS ONE* **7**, e37893. (doi:10.1371/journal.pone.0037893)

47. Elmer T, Chaitanya K, Purwar P, Stadtfeld C. 2019 The validity of RFID badges measuring face-to-face interactions. *Behav. Res. Methods* **51**, 2120–2138. (doi:10.3758/s13428-018-1180-y)
48. Claidière N, Gullstrand J, Latouche A, Fagot J. 2017 Using automated learning devices for monkeys (ALDM) to study social networks. *Behav. Res. Methods* **49**, 24–34. (doi:10.3758/s13428-015-0686-9)
49. Gelardi V, Fagot J, Barrat A, Claidière N. 2019 Detecting social (in)stability in primates from their temporal co-presence network. *Anim. Behav.* **157**, 239–254. (doi:10.1016/j.anbehav.2019.09.011)
50. Altmann J. 1974 Observational study of behaviour: sampling methods. *Behaviour* **49**, 227–266. (doi:10.1163/156853974X00534)
51. Stehlé J *et al.* 2011 High-resolution measurements of face-to-face contact patterns in a primary school. *PLoS ONE* **6**, e23176. (doi:10.1371/journal.pone.0023176)
52. Seyfarth RM. 1977 A model of social grooming among adult female monkeys. *J. Theor. Biol.* **65**, 671–698. (doi:10.1016/0022-5193(77)90015-7)
53. Bastian M, Heymann S, Jacomy M. 2009 Gephi: an open source software for exploring and manipulating networks. In *Proc. Third Int. AAAI Conf. on Weblogs and Social Media, San Jose, CA, 17–20 May 2009*. Palo Alto, CA: AAAI Press.
54. Farine D. 2018 When to choose dynamic vs. static social network analysis. *J. Anim. Ecol.* **87**, 128–138. (doi:10.1111/1365-2656.12764)