**Additional file 1**

**Statistical analysis describing the effect of sampling conditions in variations describing mosquito abundances.**

**Methods**

Generalized linear mixed effect models (GLMM) were built within the statistical software R version 3.5.0 (2018-04-23) augmented with lme4 packages for statistical analysis. The analysis was conducted to verify variation in the abundance of vector species (*Anopheles gambiae s.l*., *Anopheles arabiensis, Anopheles coluzzii*) between trapping positions (indoors, outdoors), between houses (house A and B) and among different time points (21:00-22:00; 00:00-01:00; 03:00-04:00). The response variable of mosquito abundances were highly over-dispersed, so they were modelled using a negative binomial distribution. For each species, two models were built, including time point, trapping location, house as covariates, sampling day as random effect and: i) no interaction among variables ii) interaction between trapping location and house. The two models were compared using the Akaike information criterion (AIC) and likelihood ratio test.

**Results**

The mean abundance of *An. gambiae s.l.,* *An. arabiensis*, *An. coluzzii* was best explained by a model that included the interaction between trapping position and house (Tables S1 and S2) indicating a statistically significant higher abundance of mosquitoes outdoors in house B compared for both *An. arabiensis* and *An. coluzzii* (Figure S1).

**Table S1.**  [Generalized linear mixed models built to evaluate the role of sampling conditions in explaining distributions of species abundance](https://static-content.springer.com/esm/art%3A10.1186%2Fs12936-019-3030-5/MediaObjects/12936_2019_3030_MOESM3_ESM.docx) in study. In all models a negative binomial distribution has been chosen for the response variable, including time point, position and house as fixed effect and date as random effect. AIC: Akaike Information Criterion; d.f.: degrees of freedom; χ2: chi-square test value.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Species tested** | **Model** | **AIC** | **d.f.** | **χ2** | **p-value** |
| *Anopheles gambiae* s.l. | no interaction | 650 | 1 | 20.73 | <0.005 |
| position\*house | 669 |
| *Anopheles arabiensis* | no interaction | 465 | 1 | 15.03 | 0.0001 |
| position\*house | 478 |
| *Anopheles coluzzii* | no interaction | 493 | 1 | 19.61 | <0.005 |
| position\*house | 510 |

**Table S2**.Summary of GLMM chosen for each species tested.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Species tested** | **Parameter** | **Estimate** | **Standard error** | **z-value** | **p-value** |
| *Anopheles*  *gambiae s.l.* | intercept | 3.94 | 0.17 | 21.90 | <0.0001 |
| time point 03:00-04:00 | -0.22 | 0.15 | -1.600 | 0.11 |
| time point 21:00-22:00 | 0.04 | 0.15 | 0.33 | 0.72 |
| position OUT | -0.54 | 0.17 | -3.30 | <0.005 |
| house B | -0.51 | 0.22 | -2.23 | <0.005 |
| position OUT:house B | 1.17 | 0.24 | 5.00 | <0.005 |
| *Anopheles*  *arabiensis* | intercept | 2.75 | 0.26 | 10.12 | <0.0001 |
| time point 03:00-04:00 | -0.22 | 0.15 | -1.34 | 0.17 |
| time point 21:00-22:00 | -0.22 | 0.16 | -1.24 | 0.21 |
| position OUT | -0.44 | 0.18 | -2.26 | 0.024 |
| house B | -0.72 | 0.36 | -1.97 | 0.05 |
| position OUT:house B | 1.11 | 0.28 | 4.11 | <0.0001 |
| *Anopheles*  *coluzzii* | intercept | 2.82 | 0.14 | 18.86 | <0.0001 |
| time point 03:00-04:00 | -0.21 | 0.13 | -1.47 | 0.14 |
| time point 21:00-22:00 | 0.17 | 0.14 | -1.34 | 0.18 |
| position OUT | -0.51 | 0.15 | -3.22 | 0.001 |
| house B | -0.39 | 0.18 | -2.10 | 0.034 |
| position OUT:house B | 1.09 | 0.23 | 4.81 | <0.0001 |

**Figure S1**. Boxplot showing differences in abundances of *Anopheles gambiae* s.l., *Anopheles arabiensis* and *Anopheles coluzzii* according to the house and sampling position.





