**SUPPLEMENTARY MATERIAL FOR:**

# Quantitative survival of *Leptospira interrogans* in soil and water microcosms

Arnau Casanovas-Massana[a]#, Gabriel Ghizzi Pedra[b], Elsio A. Wunder Jr[a], Peter J. Diggle[c], Mike Begon[b] and Albert I. Ko[a,d]#

[a]Department of Epidemiology of Microbial Diseases, School of Public Health, Yale University, New Haven, Connecticut, USA.
[b]Institute of Integrative Biology, The University of Liverpool, Liverpool, United Kingdom
[c]Faculty of Health and Medicine, University of Lancaster, Lancaster, United Kingdom
[d]Centro de Pesquisas Gonçalo Moniz, Fundação Oswaldo Cruz, Ministério da Saúde, Salvador, Bahia, Brazil

**#Address correspondence to:**
Arnau Casanovas-Massana: arnau.casanovas@yale.edu
Albert I. Ko**:** albert.ko@yale.edu

## 1. SUPPLEMENTARY TABLES

**Table S1.** Candidate models with effects of the covariates on $\alpha$ and $\phi$ and Likelihood Ratio test (LR) showing the best model. The final model selected is highlighted in bold.

| | Log Likelihood | k* | df** | LR |
|---|---|---|---|---|
| Null model | -739.74 | 5 | - | - |
| Treatment | -738.08 | 7 | 2 | 0.19 |
| Species | -736.08 | 7 | 2 | 0.03 |
| Method | -716.83 | 7 | 2 | <0.0001 |
| Medium | -496.32 | 15 | 10 | <0.0001 |
| Medium | -496.32 | 15 | - | - |
| Medium+species | -458.25 | 17 | 2 | <0.0001 |
| Medium+Method | -265.09 | 17 | 2 | <0.0001 |
| Medium+Method | -265.09 | 17 | - | - |
| **Medium+Method+Species** | -201.46 | 19 | 2 | <0.0001 |
| Medium+Method+Medium*Method | -257.89 | 25 | 8 | 0.072 |
| Medium+Method+Species | -201.46 | 19 | - | - |
| Medium+Method+Species+Medium*Method | -196.59 | 27 | 8 | 0.28 |

* Number of parameters;

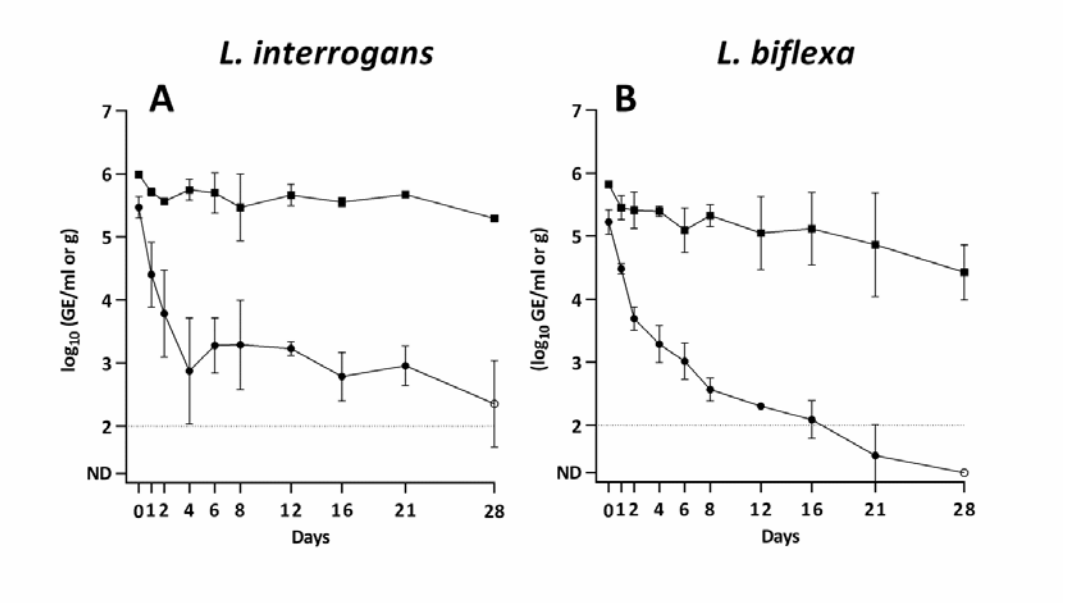
** Degrees of freedom is based in the difference between the number of parameters of each pair of comparison.

**Table S2.** Primers and probes used in this study.

| Primer or probe | Sequence (5'→ 3') | Reference |
|---|---|---|
| LipL32-45F | AAG CAT TAC CGC TTG TGG TG | (1) |
| LipL32-286R | GAA CTC CCA TTT CAG CGA TT | |
| LipL32-189P | FAM-AAA GCC AGG ACA AGC GCC G-BHQ1 | |
| rpoB-F | ATG ATG AGA CGG ATG ACT GC | This study |
| rpoB-R | CGA CGA AAC GTT TGA ACC AA | |

## 2. SUPPLEMENTARY FIGURES

**Figure S1.** Fate of *L. interrogans* (A) and *L. biflexa* (B) DNA markers measured by qPCR in microcosms of spring water (squares), soil (circles) and EMJH media (triangles) spiked with heat-killed cells. Open symbols represent data points for which at least one observation was below the limit of detection. Error bars indicate standard deviations.



## 3. SUPPLEMENTARY METHODS:

### 3.1 DNA Extraction Optimization

#### 3.1.1 Methods

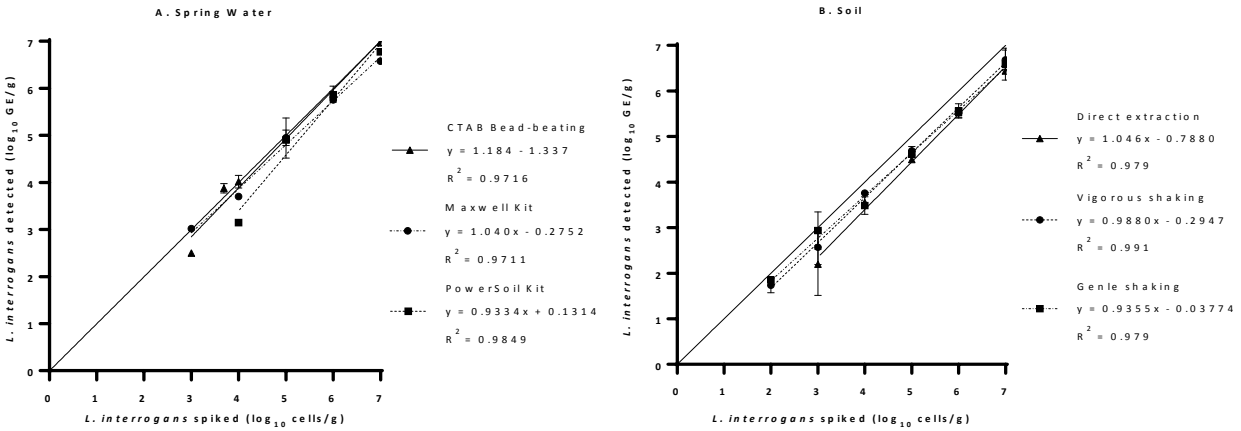
Soil samples were extracted using the Power Soil™ DNA Isolation Kit (Mobio, Carlsbad, CA, USA), with some modifications. Briefly, 9 mL of PBS was added to the 1g soil samples in a 15-mL tube, followed by vortexing for 2 min. Tubes were kept in vertical position for 20 min to allow the sediments to settle. The supernatant was recovered, centrifuged at 20,000 *g* for 20 min, the supernatant discarded and the pellet resuspended in 1 mL PBS. The sample was transferred to a 2-mL screw-cap tube included in the kit whose content (beads and lysis solution) had been previously removed. The tube was centrifuged at 20,000 *g* for 20 min and all the supernatant discarded. The beads and lysis solution were added back to the screw-cap tube and the DNA extraction protocol was followed as indicated by the manufacturer.

Spring water samples and EMJH samples were extracted using a bead beating with phenol/chloroform/isoamyl alcohol method. Briefly, 1 mL spring water sample was spiked with 100 µl of a stationary culture of *E. coli* K12 to improve cell pelleting. The mixture was centrifuged in a 2-mL screw-cap tube at 20,000 *g* for 20 min, and the supernatant carefully discarded. The pellet was mixed with 400 µL of extraction buffer

100 mM Tris-HCL pH=8.0; 1.4 M NaCl; 20 mM EDTA; 2% (w/v) CTAB; 0.2% (v/v) β-mercaptoethanol), 400 µL chlorophorm/isoamlyalcohol (24:1 v/v), 400 µl phenol, and a combination of acid-washed glass beads (1 piece of 3 mm, 0.25 g of 1 mm and 0.37g of 0.1-0.11 mm). Then, the tube was bead beat for 5 min using a Bullet Blender (Next Advance, NY, USA), followed by centrifugation at 14,000 *g* for 10 min. The aqueous layer was transferred to a new 1.5-mL microtube containing 500 µL of chlorophorm/isoamlyalcohol (24:1 v/v), vortexed for 2 min and centrifuged at 14,000 *g* for 10 min. The aqueous layer was transferred to a new tube, and the DNA precipitated using 1 volume of isopropanol and incubating the sample overnight at room temperature. After incubation, the samples were centrifuged at 14,000 *g* for 15 min, and the DNA pellet was double washed with 70% ice-cold ethanol. The pellet was allowed to dry out, and finally resuspended in 50 µL of Tris-HCl 10mM pH=8.0.

### 3.1.2 Results

**Figure 1.** Modelling of the DNA extraction methods evaluated for spring water (A) and soil matrices (B).

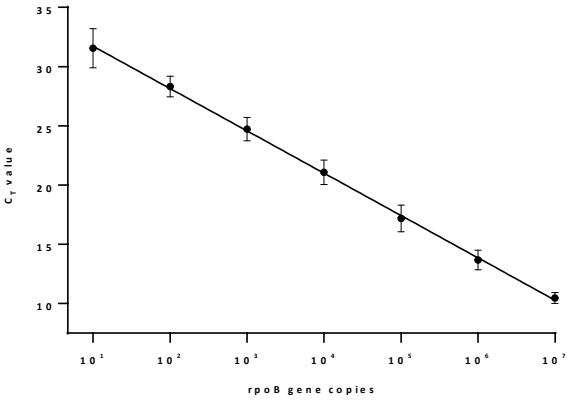



### 3.2 qPCR assays

#### 3.2.1 Design of a qPCR for *L. biflexa* Patoc

*L. biflexa* Patoc was quantified using a newly designed SYBR-Green® reaction amplify the *rpoB* gene. Specific primers were designed to amplify the rpoB gene using the complete genome (access number: NC_010602). After primer design, a BLAST analysis was performed to confirm their specificity for the target genes. Reaction mixtures (20 µl) contained 10.0 µl IG™ SYBR-Green® Supermix (BioRad), 300 nM of each primer, 0.2 µg/µL of bovine serum albumin (Invitrogene) and 5 µl of DNA extract. The thermal-cycler conditions were as follow: an initial setup of 2 min at 50°C, followed by 10 min at 95°C, and 40 cycles of amplification (15 s at 95°C and 1 min at 60°C). Product specificity was confirmed by melting curve analysis (60-95°C, 0.3°C per read, 30 s hold). Standard curves were constructed using genomic DNA obtained from strain Patoc1 with concentrations ranging from $10^7$ to $10^0$ GE (genomic equivalents)/5µL, based on its respective genome size (2). All the samples and standards were run in duplicate and non-template controls were included in each row of the plates. All amplifications were performed in a 7500 Fast Real-Time PCR System (Life Technologies). Pair primers used for are detailed in Table S2. Calibration curves were included in each qPCR run and efficiencies were always higher than 90.5%.

**Figure 2.** *L. biflexa* qPCR standard curve prepared with plasmids as described above for *rpb*O gene. The data presented are the averages of six replicates performed in six independent experiments.




### 3.2.2 *L. interrogans* qPCR assay

*L. interrogans* was quantified using a TaqMan® assay targeting the *lipL32* gene with minor modifications (3). Reaction mixtures (25 µL) contained 12.5 µL Platinum® qPCR SuperMix (Life Technologies), 500 nM of primers LipL32-45F and LipL32-286R (Table S2), 100 nM of LipL32-189P probe, 0.2 µg/µL of bovine serum albumin (Ambion) and 5 µL of DNA template. Amplifications were performed in a 7500 Fast Real-Time PCR System (Applied Biosystems). The thermal cycling conditions consisted of an initial setup of 2 min at 50°C, followed by 2 min at 95°C, and 40 cycles of amplification (15 s at 95°C and 1 min at 60°C). All samples, controls and calibrators were run in duplicate. Non-template controls were randomly included in each row of the 96-well plates. Genomic DNA from *L. interrogans* serovar Copenhageni strain Fiocruz L1-130 was extracted using the automated Maxwell® 16 Cell DNA Purification Kit (Promega) and quantified with Nanodrop 2000 (Thermo Fisher Scientific). The DNA was serially 10-fold diluted to construct calibration curves with concentrations ranging from $2 \times 10^9$ to $2 \times 10^2$ GEq/mL. Genomic equivalents were calculated based on the genome size of the Fiocruz L1-130 strain (4) . Calibration curves were included in each qPCR run and efficiencies were always higher than 92.5%.

### 3.2.3 Internal Amplification Control (IAC)

qPCR inhibition was monitored in each environmental sample by qPCR using the an IAC plasmid (3). Three DNA extracts from each microcosms replicate were randomly selected and spiked with 200 copies of the IAC in duplicate. As criteria for inhibition, we established a limit of 2 standard deviations above the mean quantification cycle ($C_q$ of 32.63 + 1.51), which was based on $C_q$ values obtained from 30 control amplifications of IAC in ultrapure water in 7 independent qPCR runs. DNA extracts were assumed to be free of qPCR inhibitors when their mean $C_q$ value fell below that limit. There was no evidence of inhibition of the molecular assays in any of the microcosms.

## 3.3 Optimization of a PMA treatment to discriminate live and dead *L. interrogans* cells

### 3.3.1 Methods

The ability of propidium monoazide (PMA) to selectively amplify DNA from membrane-intact L. interrogans cells in spring water and soil was investigated. A mid log phase culture of L. interrogans Copenhageni strain Fiocruz L1-130 was adjusted to a concentration of 108 cells/mL in a total volume of 5 mL of EMJH medium and then divided into 2.5 mL aliquots. To disrupt the membranes, one of the aliquots was heat-killed by exposure to 80°C for 15 min using a water bath with subsequent cooling to room temperature for 20 min.

For spring water experiments 100 µL of each culture were spiked to 900 µL of spring water in light-transparent 1.5 mL microtubes to achieve a final concentration of 107 cells/mL. A solution of 2 mM of PMA (Biotium) was prepared with ultrapure water and added to the 1 mL spring water cell suspensions to obtain a final concentration of 10µM. Samples were thoroughly mixed by inverting the tubes several times after the PMA addition and incubated in the dark for 15, 30, 45 or 60 min with frequent mixing. After the PMA cross-linking time, tubes were placed on ice to avoid excessive heating and exposed for 5 min to a 600W halogen light source placed 20 cm above the samples. Then, 100 µL of a stationary culture of E. coli K12 was added to the tubes to improve cell pelleting and were centrifuged at 20,000 g for 20 min. The supernatant was discarded, and the pellet recovered in 400 µL of PBS by vortexing for 1 min. Finally, DNA was extracted using the automated Maxwell® 16 Cell DNA Purification Kit (Promega) and quantified using the qPCR procedure described above.

For soil samples, 100 µL of each culture were spiked to 0.9 g of soil in 15 mL tubes to achieve a final concentration of 107 cells/g. Then, 9 mL of PBS was added to the samples, followed by mixing for 1 hour in a horizontal agitator at maximum speed. Tubes were kept in vertical position for 20 min to allow the sediments to settle. The supernatants were recovered, centrifuged at 8,000 g for 20 min, and the pellet resuspended in 1 mL PBS. The samples were transferred to light-transparent 1.5 mL microtubes and PMA was added at different final concentrations (50 µM, 100 µM and 200 µM). Samples were thoroughly mixed by inverting the tubes several times after the PMA addition and incubated in the dark for 15 minutes. Then, tubes were placed on ice and exposed for 15 min to a 600W halogen light source placed 20 cm above the samples. After the light exposure, soil samples were extracted following the procedure detailed in the Supplementary Material 1, and quantified using the qPCR procedure described above.

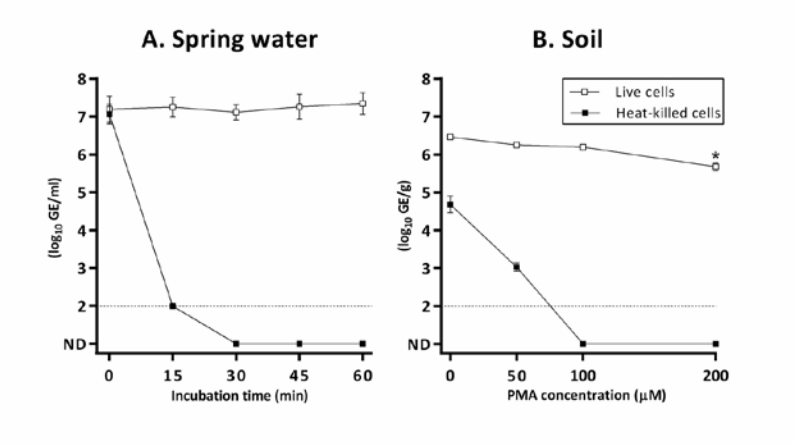
All manipulations of PMA were performed under minimal light to avoid any chemical damage to the light-sensitive PMA structure and all experiments were conducted in three biological replicates. A microtube with live cells and another with heat-killed not treated with PMA were included as a control.

### 3.3.2 Results

Increasing incubation time of PMA with samples in spring water lead to a decrease of the qPCR ability to detect DNA from heat-killed cells without affecting the detection of live cells (Fig. 1 A). As a result, we selected a 60 min treatment with 5 µM PMA for subsequent experiments. In Brazilian soil, the increase of PMA concentration also decreased the qPCR signal in heat-killed cells. However, the treatment with 200 µM of PMA decreased significantly the detection from live cells (Fig. 1B). We selected a 15 min treatment with 100 µM of PMA for subsequent experiments, since it showed an efficient elimination of heat-killed cells signal without compromising the live cells. Unfortunately, after testing various PMA concentrations, and incubation and illumination times (data not shown) the attempts to optimize a PMA treatment for US soil were unsuccessful.

**Figure 3.** Effect of the PMA treatment on the DNA quantification by qPCR. (A) Effect of the increasing incubation time with PMA on the signal detected in spring water (B) Effect of the

increasing PMA concentration on the signal detected in Brazilian soil. White squares denote live cells and black squares heat-killed cells.




### 3.4 Modeling

#### 3.4.1 Model

Following Peleg and Cole, 1988 and van Boekel, 2002 (5, 6) we used a Weibull distribution to model the survival time, *T*, with survivor function:

$$S(t;\phi,k) = P(T > t) = \exp\left(-\left({}^{t}/_{\phi}\right)^{k}\right) : t \geq 0 \quad (1)$$

The parameter $k$ determines the shape of the survival curve, whilst $\phi$ defines how stretched the shape is; specifically, $\phi$ is the expectation (average value) of $T_k$. it follows from (1) that if we consider a closed population of cells with initial concentration $\mu_0$ at time *t=0* and measure the concentration of surviving cells at a subsequent time *t*, the expected concentration at time t is $\mu_t = \mu_0 S(t;\phi,\kappa)$. However, in our experiments, we observed that a proportion of the cells appeared to survive well beyond the maximum follow-up time. We therefore extended the model to $\mu_t = \mu_0(\alpha + (1-\alpha)S(t;\phi,\kappa))$, where $\alpha$ is the proportion of long-term survivors.

We now consider a set of experiments, *i=1,..,r*, the *ith* of which is characterized by the values of a set of covariates $x_i$, In each experiment we measure the concentration at a sequence of times $t_j : j = 1,2,\ldots,m$. Our model for the complete set of experiments becomes:

$$\mu_{ij} = \mu_0 * (\alpha_i + (1-\alpha_i) * \left(S(t_j;\phi_i,\kappa)\right) \quad (2)$$

In equation (2), the effects of the covariates on the values of $\phi$ and $\alpha$ were explored to determine if there were any differences in survival between species, treatment and method of quantification by specifying log-linear and logistic models for $\phi$ and $\alpha$ respectively, hence:

$$\phi_i = e^{x_i'\beta} \tag{3}$$

and

$$\alpha_i = {}^{1}/_{(1 + e^{-x_i'\gamma})} \tag{4}$$

Finally, we assume that observed concentration $Y_{ij}$ are independent and Normally distributed,

$$Y_{ij} = \mu_{ij} + Z_{ij} \quad i = 1, \dots, r; j = 1, \dots, m \tag{5}$$

where $Z_{ij}$, the observation level residuals, are Normally distributed $Z_{ij} \sim N(0, \tau^2)$.

### 3.4.2 Log-likelihood

The log-likelihood for the complete set of data contains contributions of two kinds: measured values $y_{ij}$ and results recorded only as below-detection, representing values $y_{ij} < d$. Let $f(y; \mu, \tau^2)$ denote the probability density, and $F(y; \mu, \tau^2)$ the cumulative probability distribution, of the Normal distribution mean $\mu$ and variance $\tau^2$. Then, the log-likelihood for the complete set of parameters $\theta = (\beta, \gamma, \mu_0, \tau^2)$ is:

$$l(\theta) = \sum_{i=1}^{r} \sum_{j=1}^{m} l_{ij}(\theta) \tag{6}$$

where:

for observations $y_{ij}$ , $l_{ij}(\theta) = \log f(y; \mu_{ij}, \tau^2)$;

for observations $y_{ij} < d$ , $l_{ij}(\theta) = \log F(d; \mu_{ij}, \tau^2)$;

The parameters were estimated by optimizing the log-likelihood function, using the *optim()* function in the R software (7)

### 3.4.3 Confidence intervals

95% Confidence intervals for individual parameters $(\beta, \lambda, \mu_0, k)$ were calculated as

$$\hat{\theta} \pm 1.96SE(\hat{\theta}) \tag{7}$$

where SE denotes the square root of the variance of $\hat{\theta}$ as given by the information matrix. To calculate 95% confidence intervals for $\phi_i$ we calculate the variance of $\log(\phi)$ as $v = x_i' Var(\hat{\beta}) x_i$, calculate limits $a$ and $b$ as $x_i'\hat{\beta} \pm 1.96\sqrt{v}$, then transform $a$ and $b$ to give the confidence interval $(e^a, e^b)$. Similarly, to calculate 95% confidence intervals for $\alpha_i$, we calculate $v = x_i' Var(\hat{\gamma}) x_i$, calculate limits $a$ and $b$ as $x_i'\hat{\gamma} \pm 1.96\sqrt{v}$, then transform $a$ and $b$ to give the confidence interval $\left({}^{1}/_{(1+e^{-a})}, {}^{1}/_{(1+e^{-b})}\right)$.

### 3.4.4 Checking assumptions

A plot of standardized residuals against fitted values was inspected to check the fit of the model to the data. The plot (Figure 1) indicates a reasonably good fit. NOTE: only one panel is needed.

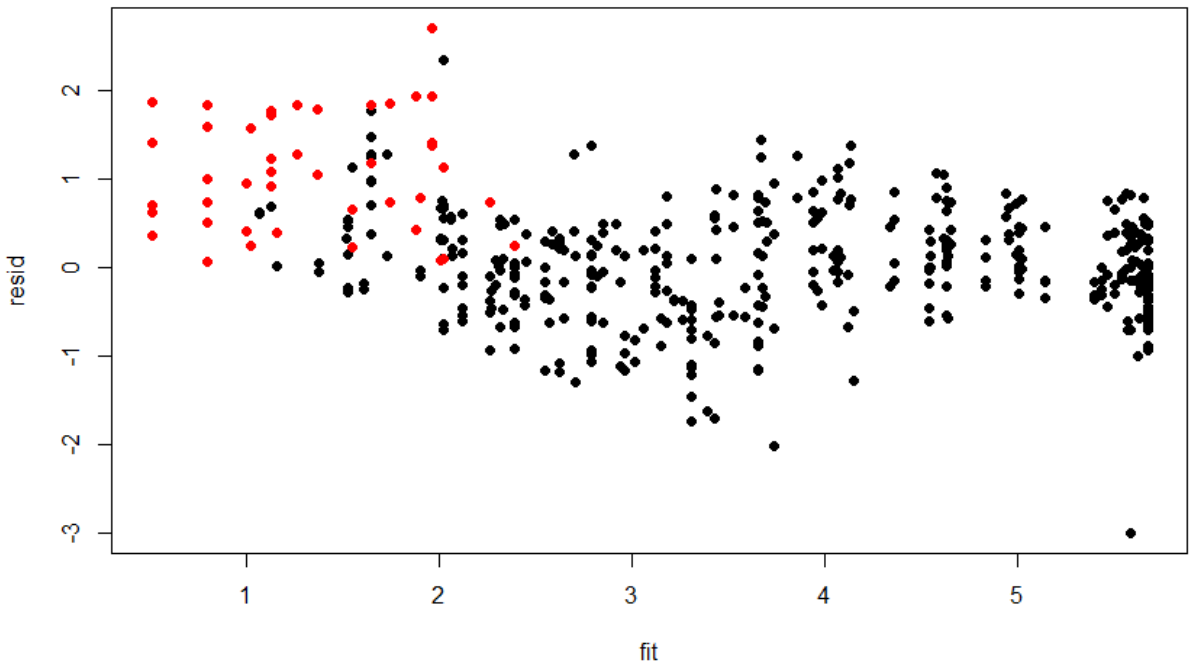



**Figure:** Standardized residuals vs fitted values. The red dots are the fitted values for below detection limit.

### 3.4.5 Model selection

The selection of the covariates in the model was based on a Likelihood ratio test (LRT), whereby twice the difference between the log-likelihoods of two nested models is compared with critical values of a chi-squared distribution with degree of freedom equal to the difference in the number of parameters in the two models. Firstly, the main effect of each covariate was tested against a null model with no covariates. Then, interactions were tested against the selected main-effects-only model.

## 4. REFERENCES


1. **Stoddard RA**, **Gee JE**, **Wilkins PP**, **McCaustland K**, **Hoffmaster AR**. 2009. Detection of pathogenic Leptospira spp. through TaqMan polymerase chain reaction targeting the LipL32 gene. Diagn. Microbiol. Infect. Dis. **64**:247–55.
2. **Picardeau M**, **Bulach DM**, **Bouchier C**, **Zuerner RL**, **Zidane N**, **Wilson PJ**, **Creno S**, **Kuczek ES**, **Bommezzadri S**, **Davis JC**, **McGrath A**, **Johnson MJ**, **Boursaux-Eude C**, **Seemann T**, **Rouy Z**, **Coppel RL**, **Rood JI**, **Lajus A**, **Davies JK**, **Médigue C**, **Adler B**. 2008. Genome sequence of the saprophyte Leptospira biflexa provides insights into the evolution of Leptospira and the pathogenesis of leptospirosis. PLoS One **3**:e1607.
3. **Casanovas-Massana A**, **Costa F**, **Riediger IN**, **Cunha M**, **de Oliveira D**, **Mota DC**, **Sousa E**, **Querino VA**, **Nery N**, **Reis MG**, **Wunder EA**, **Diggle PJ**, **Ko AI**. 2018. Spatial and temporal dynamics of pathogenic Leptospira in surface waters from the urban slum environment. Water Res. **130**:176–184.
4. **Nascimento ALTO**, **Verjovski-Almeida S**, **Van Sluys MA**, **Monteiro-Vitorello CB**, **Camargo LEA**, **Digiampietri LA**, **Harstkeerl RA**, **Ho PL**, **Marques MV**, **Oliveira MC**, **Setubal JC**, **Haake DA**, **Martins EAL**. 2004. Genome features of Leptospira interrogans serovar Copenhageni. Brazilian J. Med. Biol. Res. **37**:459–478.
5. **Peleg M**, **Cole MB**. 1998. Reinterpretation of Microbial Survival Curves. Crit. Rev. Food Sci. Nutr. **38**:353–380.
6. **van Boekel MAJ.** 2002. On the use of the Weibull model to describe thermal inactivation of microbial vegetative cells. Int. J. Food Microbiol. **74**:139–159.
7. **R Core Team**. 2013. A language and environment for statistical computing. 3.1. R Foundation for Statistical Computing, Vienna, Austria.