

Supplementary Material for

Rousset, François and Ferdy, Jean-Baptiste. Testing environmental and genetic effects in the presence of spatial autocorrelation — *Ecography* XX: xxx–xxx.

Appendix A Toy example of bootstrap correction

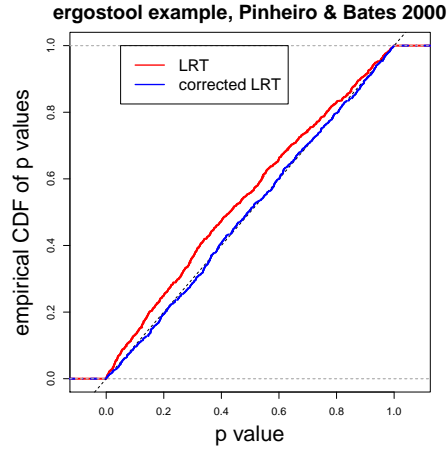


Figure A1: Comparison of LR tests and bootstrap-corrected LR tests for the simulation design considered by Pinheiro and Bates (2000, p. 88)

The full model is $y_{ij} = \beta_j + b_i + e_{ij}$ for $j = 1, \dots, 4$ and $i = 1, \dots, 9$ and the hypothesis tested is that all β_j s are equal. 1000 samples were simulated under the null hypothesis with $\text{var}(b_i) = 0.8264$ and $\text{var}(e_{ij}) = 4.083$. 200 bootstrap replicates were computed for each sample.

Appendix B Choices among fitting methods and algorithms

In this Appendix we describe possible alternatives for fitting methods and algorithms and explain the choices further made among them in this work.

B.1 REML or not REML

In a linear mixed model, the ML estimators of the dispersion parameters are known to be biased. This has led to the definition of restricted ML (REML) for linear mixed models, wherein the dispersion parameters are estimated by maximizing a conditional log-likelihood given fixed effect estimates, $\ell(y; \boldsymbol{\beta}, \lambda, \phi) - \ell(\hat{\boldsymbol{\beta}}; \boldsymbol{\beta}, \lambda, \phi)$, rather than the log-likelihood $\ell(y; \boldsymbol{\beta}, \lambda, \phi)$. However, likelihoods from REML fits are not directly suitable for likelihood ratio tests of fixed effects (e.g., Pinheiro and Bates, 2000), restricting the utility of the REML approach. Some workarounds still using a form of REML correction are nevertheless possible, as shown for LMMs by Welham and Thompson (1997). We implemented similar modifications of REML for GLMMs as further detailed in Appendix E.

B.2 Fitting algorithms

In non-spatial models each step of the iterative procedure can be performed through efficient *ad hoc* algorithms, such as iteratively reweighted least squares (IRLS), to fit the fixed effect parameters given the dispersion parameters. However, efficient *ad hoc* algorithms may not be available to estimate spatial correlation parameters given other parameters. This led us to consider more generic maximization methods for estimation of correlation parameters, and then as a viable method for joint estimation of correlation and dispersion parameters. IRLS is still always used to estimate fixed effect parameters given other parameters. Depending on the objective function used in numerical maximizations (h , p_v or $p_{\beta,v}$), and whether iterative procedures are used to estimate dispersion parameters given fixed effect parameters, or correlation parameters given all other parameters, one can then define many variants of the joint estimation procedure, yielding different joint fits and LR tests.

Negative log-LR are easily obtained if the likelihood is imperfectly maximized in the full model, so that such imperfect fits are more easily detected when likelihood ratio tests are assessed than when only bias and variance of estimators are

evaluated. We found it easier and faster to maximize the likelihood using generic maximization for estimation of correlation parameters. Thus, although our procedures can fit models by either iterative or generic maximization methods, the results presented below were obtained by generic maximization, unless noted otherwise. We compared three main cases, for the default simulation design described below.

First, all parameters were estimated by maximization of p_v . This amounts to maximum likelihood estimation of all parameters, exactly in LMMs, approximately in other GLMMs, and will be described as “ML”, for short. This also ensures that the log-likelihood ratio (the difference in p_v values among the models compared) is positive, and is thus suitable for testing fixed effects. This is the default method considered in this work. For binary data, we also considered ML fits using the second-order approximation to the likelihood ($p_v^s(h)$ in Noh and Lee, 2007), but as explained later this brings no improvement in our application, and will not be detailed.

Second, we considered a partial REML procedure. However, the main difference of the partial REML tests with those derived from the ML fits was a distortion of the distribution of p -values, observed in most cases for high (> 0.5) p -values, and less systematically for low p -values. For this reason, the REML procedure and its performance are detailed only in Appendix E and are not further discussed.

Finally, for Poisson and binomial GLMMs (including the binary case), we considered “PQL/L”, where β is estimated by maximization of h as in standard PQL, and all dispersion and correlation parameters are estimated by generic numerical maximization of p_v . REML estimation of dispersion, as described above, was also considered but one simulation showed that it again mainly resulted in a distortion of the distribution of p -values for high p -values. Thus, it is not presented in this paper, although it also appeared to result in better λ estimation, as expected. The present PQL/L method is therefore similar to the “HL(0,1)” method (Noh and Lee, 2007) except that it does not use an REML correction. The original PQL of Breslow and Clayton (1993) further differs from the HL(0,1) method as it neglects the derivatives, with respect to the estimated dispersion parameters, of the GLM weights in the Hessian $H(h, \beta, \mathbf{v})$ (Lee and Nelder, 2001; Lee et al., 2006, p. 189). In this respect, PQL/L is similar to HL(0,1) and different from the original PQL, as numerical maximization of p_v implicitly takes into account the derivatives of the GLM weights in $H(h, \mathbf{v})$.

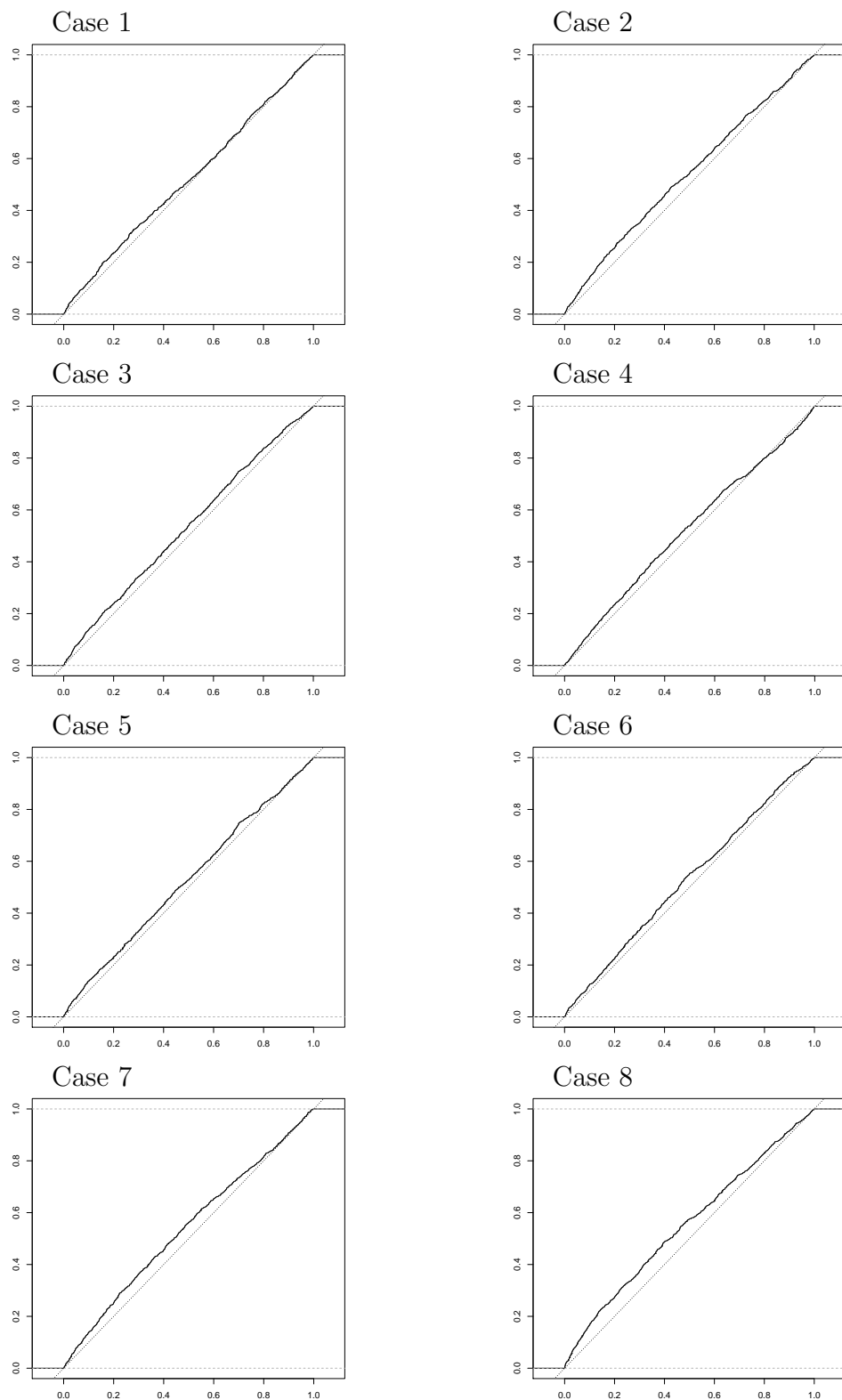
In summary, we will discuss ML and PQL/L fits, which are obtained by maximizing p_v over dispersion and correlation parameters, where for each given values of these parameters the fixed effects are estimated either by maximization of p_v or of h ,

using IRLS algorithms.

Appendix C ML fits

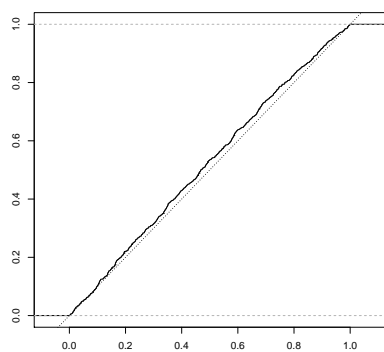
Figure C1: Distributions of p -values of ML-based tests
All plots are empirical CDFs of p -values, numbered as successive rows in Table 1.

Linear mixed model

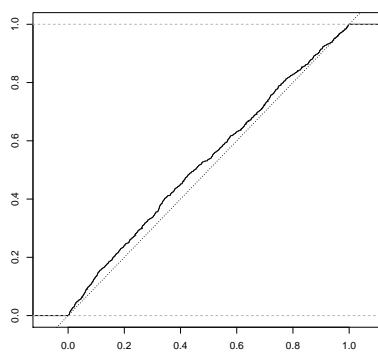


Poisson GLMM

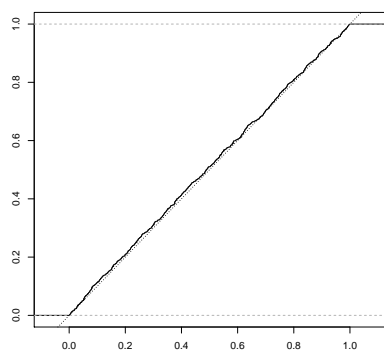
Case 1



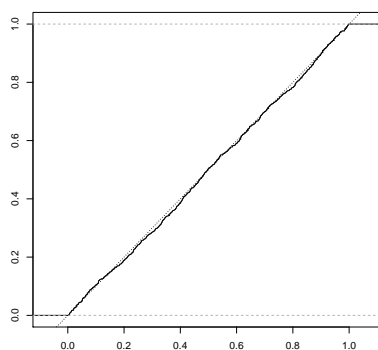
Case 2



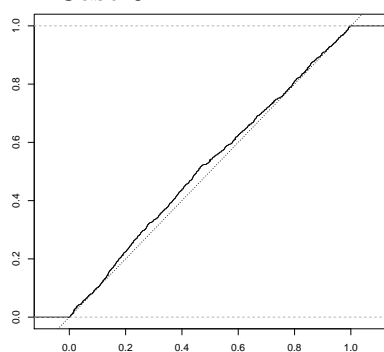
Case 3



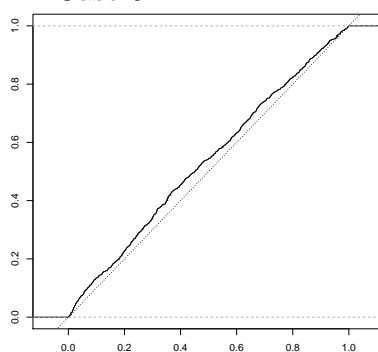
Case 4



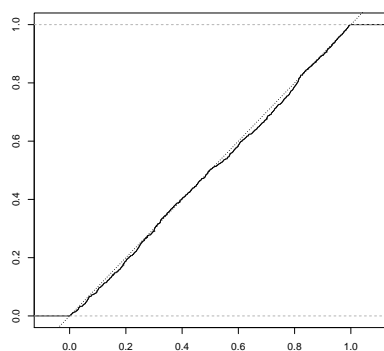
Case 5



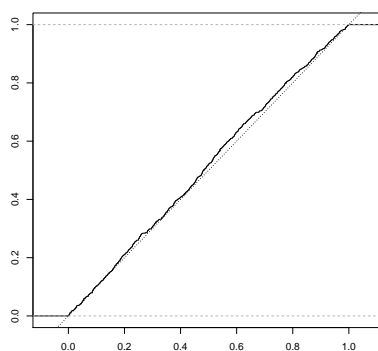
Case 6



Case 7

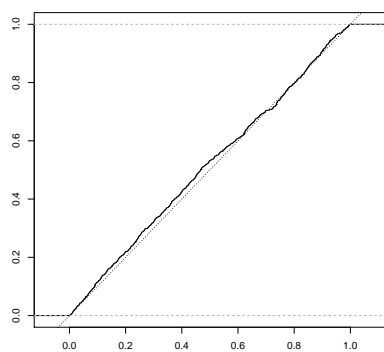


Case 8

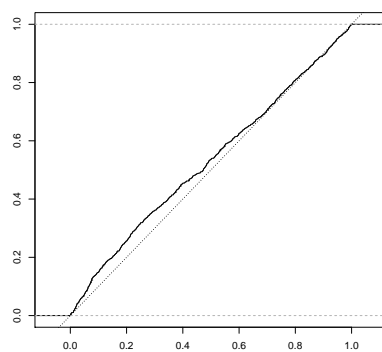


Binomial GLMM

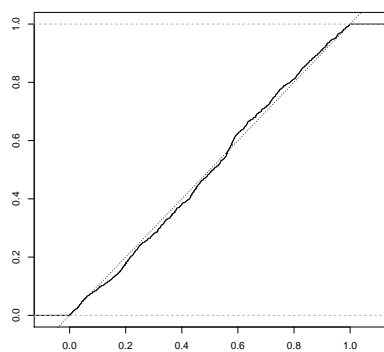
Case 1



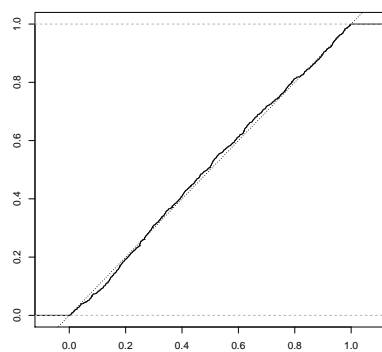
Case 2



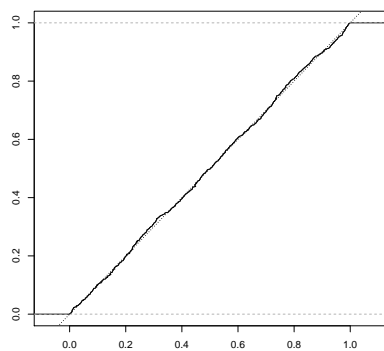
Case 3



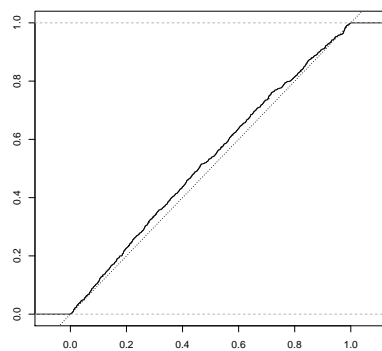
Case 4



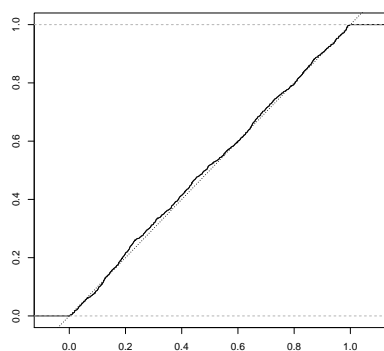
Case 5



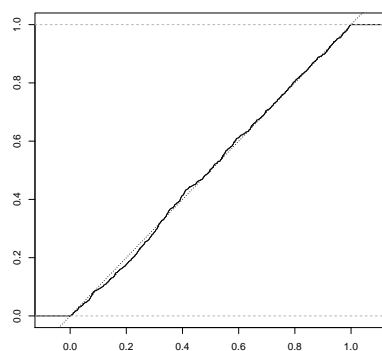
Case 6



Case 7

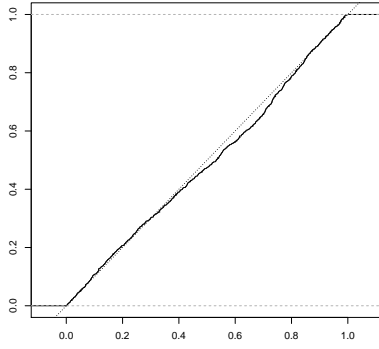


Case 8

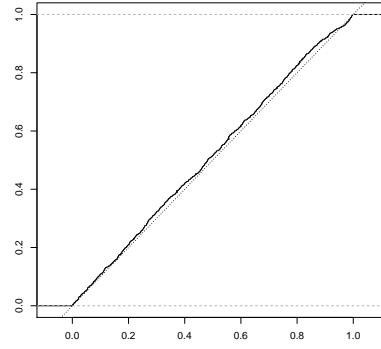


Binary GLMM with $\hat{\lambda} < 5$

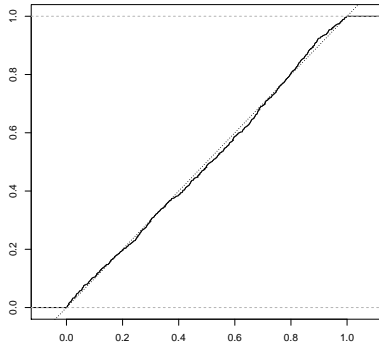
Case 1



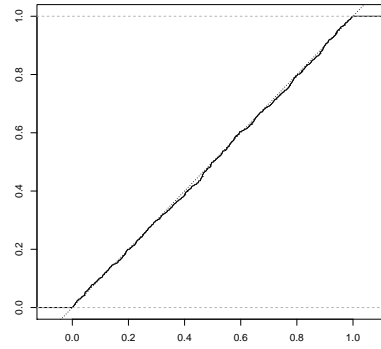
Case 2



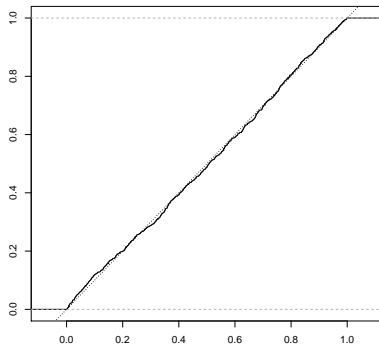
Case 3



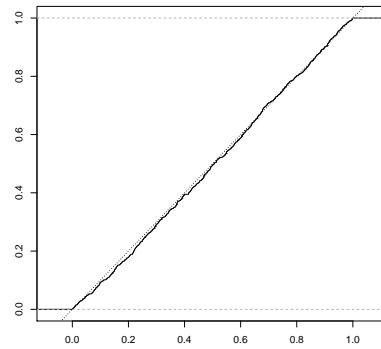
Case 4



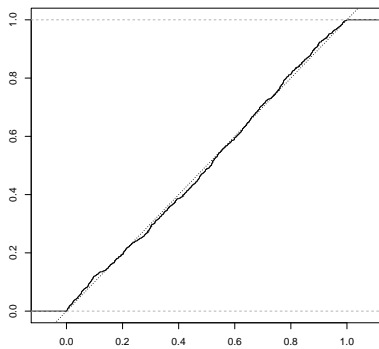
Case 5



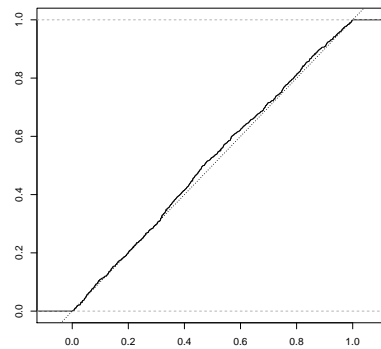
Case 6



Case 7



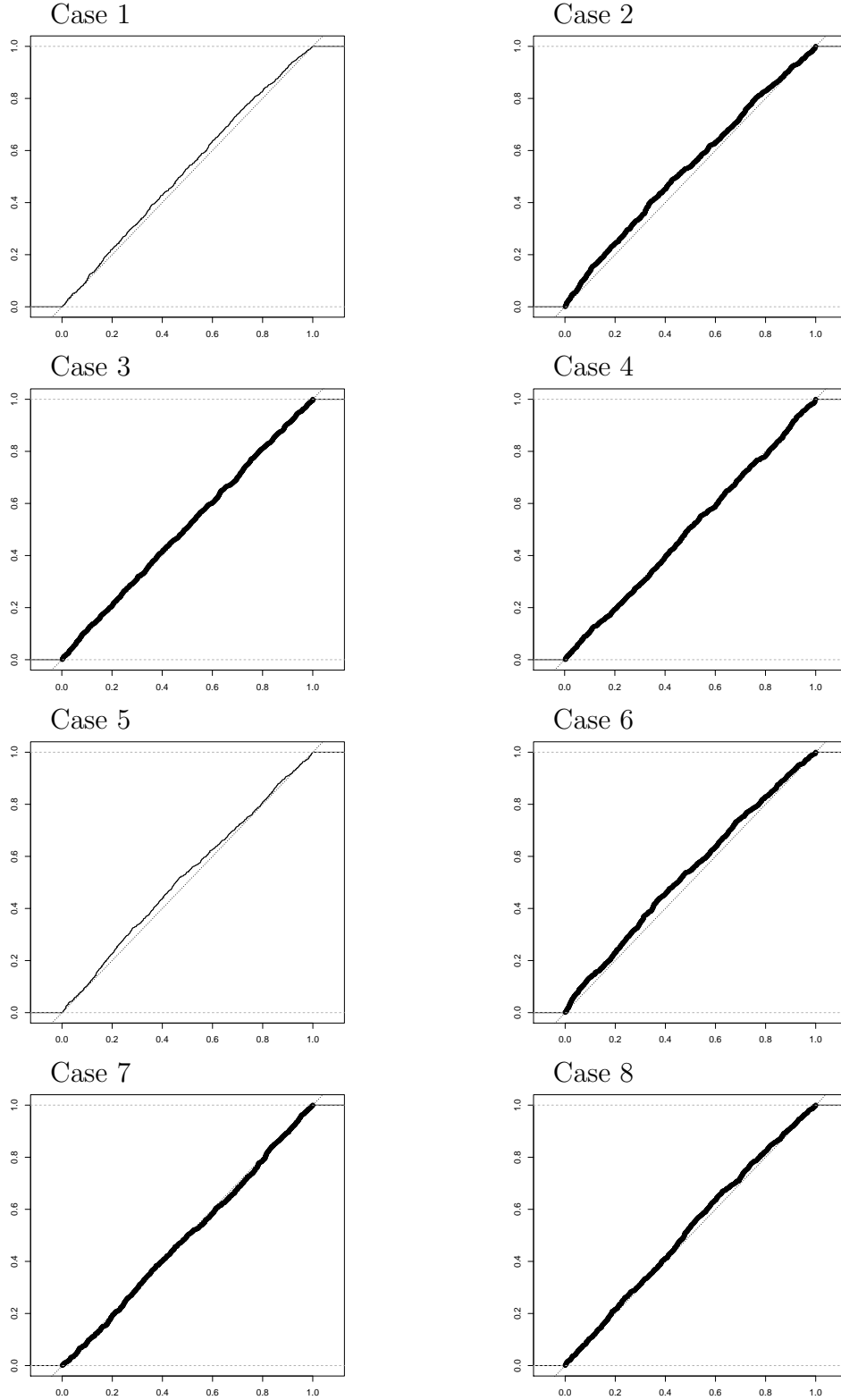
Case 8



Appendix D PQL/L fits

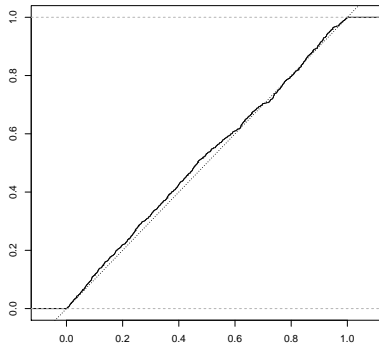
Figure D1: Distributions of p -values of PQL/L-based tests
All plots are empirical CDFs of p -values, numbered as successive rows in Table 2.

Poisson GLMM

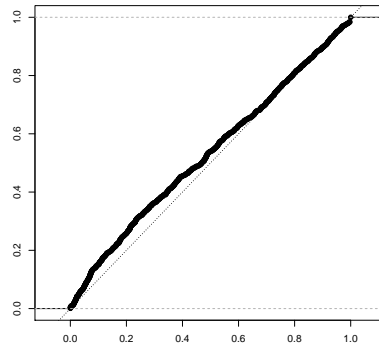


Binomial GLMM

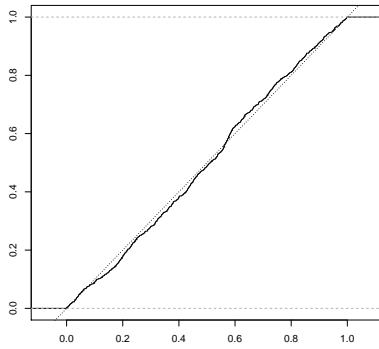
Case 1



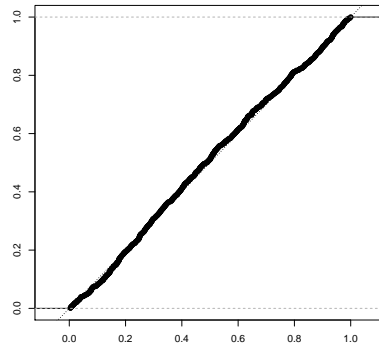
Case 2



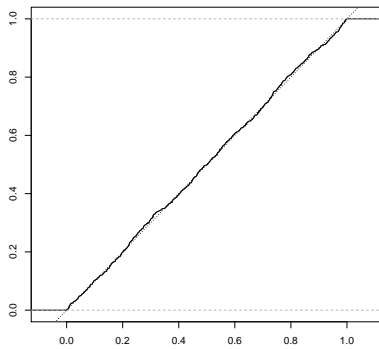
Case 3



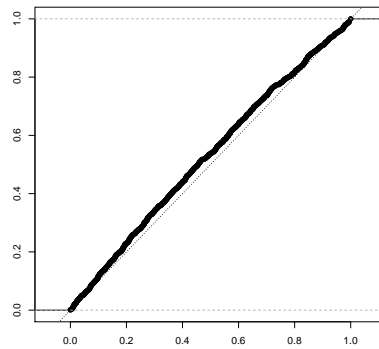
Case 4



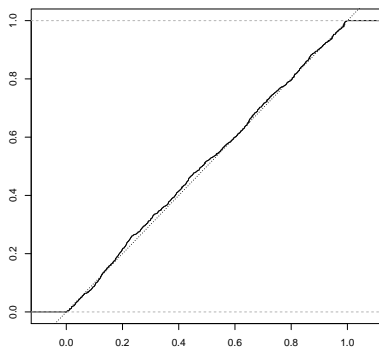
Case 5



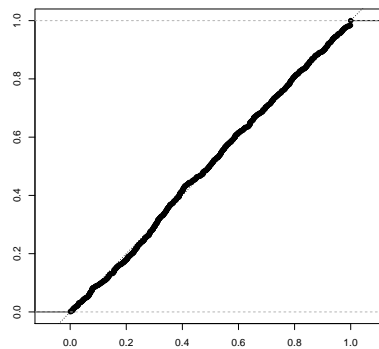
Case 6



Case 7

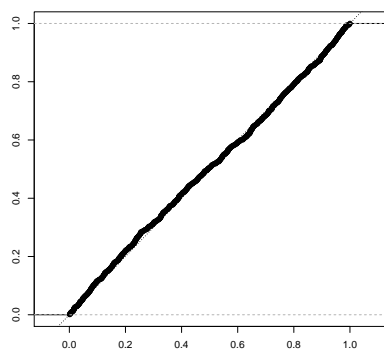


Case 8

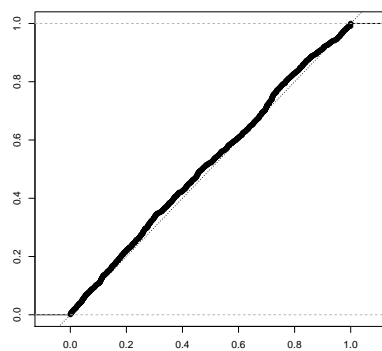


Binary GLMM

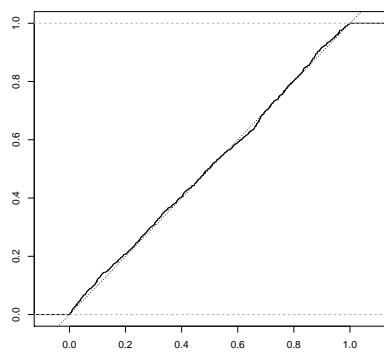
Case 1



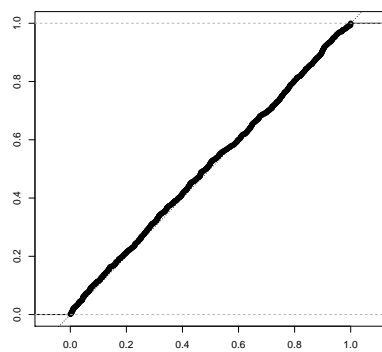
Case 2



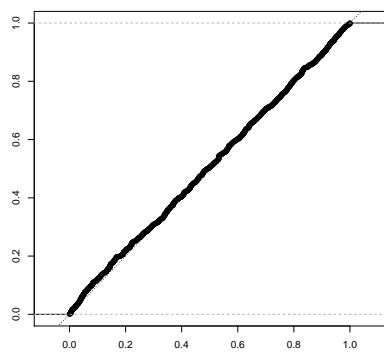
Case 3



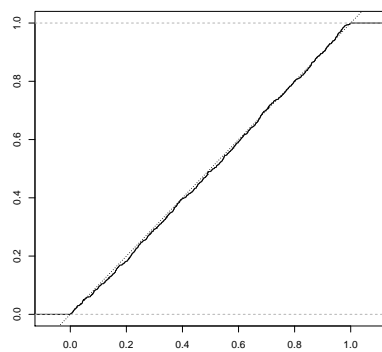
Case 4



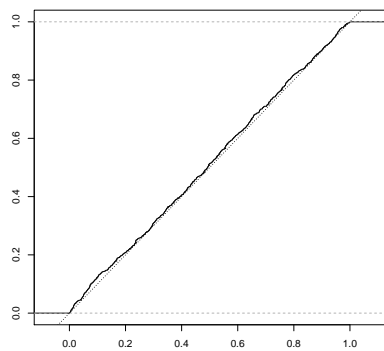
Case 5



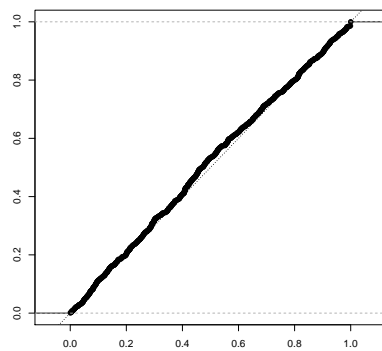
Case 6



Case 7



Case 8



Appendix E REML fits

Both PQL and Lee and Nelder’s (2001) formulation of more general Laplace approximation methods provide an extension of REML for GLMMs, in the form of the maximization of an approximation $p_{\beta,v}(h)$ of the conditional log-likelihood given the fixed effect estimates. Here

$$p_{\beta,v}(h) = h - \frac{1}{2} \log |H(h, \beta, \mathbf{v}) / (2\pi)| \quad (1)$$

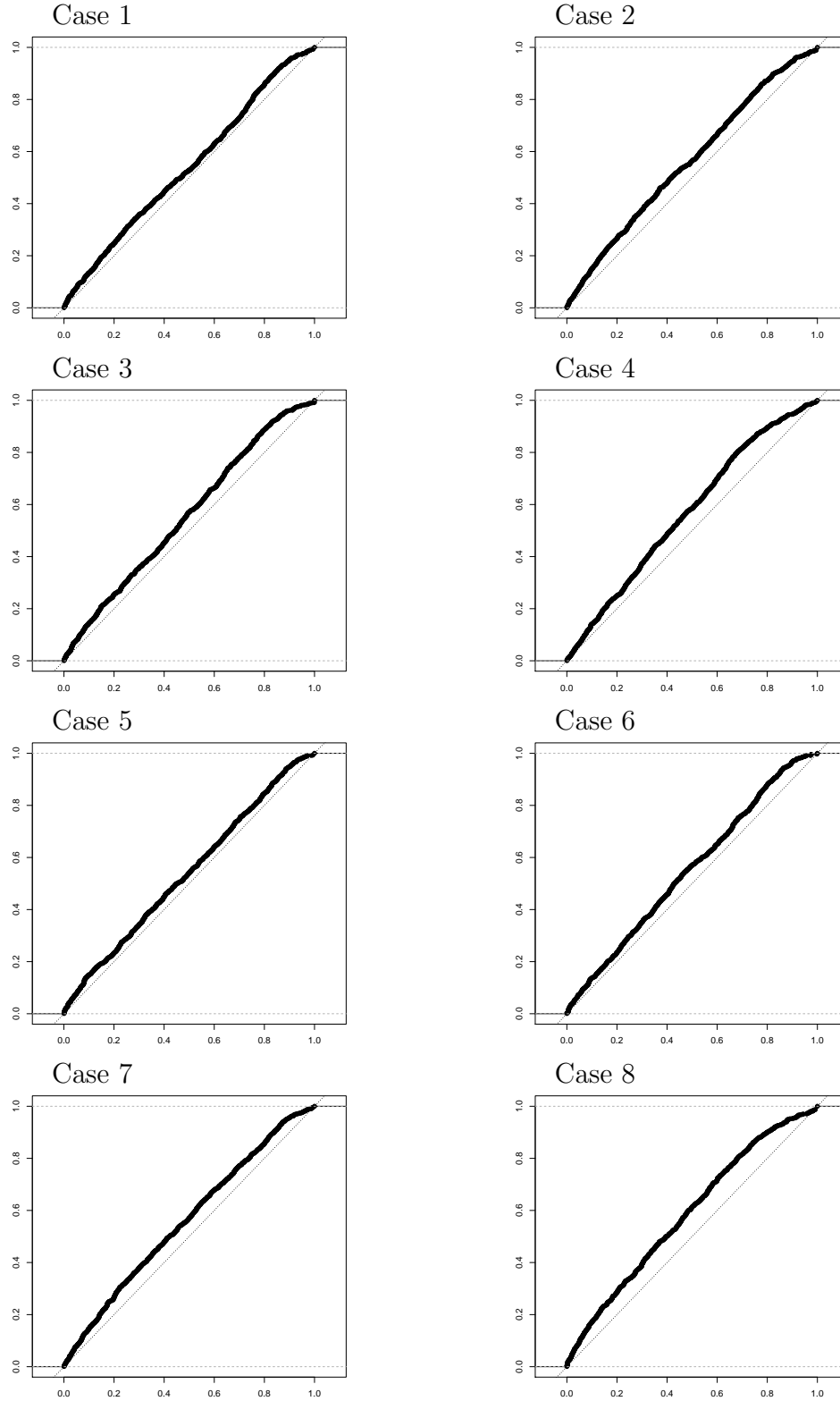
where $H(h, \beta, \mathbf{v})$ is the Hessian matrix of the h likelihood with respect to both the fixed effects β and the random effects \mathbf{v} . Compared to $H(h, \mathbf{v})$, $H(h, \beta, \mathbf{v}) \equiv H(h, \mathbf{X}, \beta, \mathbf{v})$ additionally depends on the design matrix \mathbf{X} for the fixed effects (e.g., Lee and Nelder, 1996, p. 631) and thus differs between “null” and “full” models with different fixed effects.

Likelihoods from REML fits are not directly suitable for likelihood ratio tests of fixed effects (e.g., Pinheiro and Bates, 2000), restricting the utility of the REML approach. However, some workarounds still using a form of REML correction are possible: in linear mixed models, one can compare modified restricted likelihoods of the full and null models where the same design matrix (for example that of the full model, \mathbf{X}_{full}) is used in the computation of the Hessian matrix for both models (Welham and Thompson, 1997). In GLMMs, we likewise considered the modified approximations to the restricted likelihood, $h - \frac{1}{2} \log |H(h, \mathbf{X}_{\text{full}}, \beta, \mathbf{v}) / (2\pi)|$. For given correlation parameters, fixed effects and dispersion parameters were estimated by the iterative procedure alternating $p_v(h)$ and $p_{\beta,v}(h)$ maximization, and correlation parameters were estimated by maximizing the attained p_v value. Following the modification of REML described above, \mathbf{X}_{full} was used in the computation of the Hessian for the null model.

Estimation of fixed effect parameters by maximization of $p_v(h)$ depends on random effect parameters, and conversely maximization of $p_{\beta,v}(h)$ depends on fixed effect estimates. Joint estimation can then be performed by an iterative procedure alternating these two steps, which can further be extended to estimate correlation parameters. However, in GLMMs the joint estimation of all parameters using both p_v and $p_{\beta,v}$ (or modified $p_{\beta,v}$) cannot be exactly formulated as the maximization of a single function. This has two consequences. First, this means that different joint estimates may be reached by this iterative procedure or by a procedure that maximizes a profile p_v , where the profile is over all dispersion and correlation parameters, whether estimated by ML or REML. Second, there is no certainty that the p_v of a constrained model will be lower than the p_v of the unconstrained model. Then, a negative log-LR can be obtained. Asymptotic arguments (parameter orthogonality, e.g. Lee et al., 2006, p.35) suggest that negative log-LR could be a relatively rare occurrence, and be small in absolute magnitude. This expectation is borne out in simulations, but negative log-LR do occur. For similar reasons, negative log-LR can also occur in PQL/L fits.

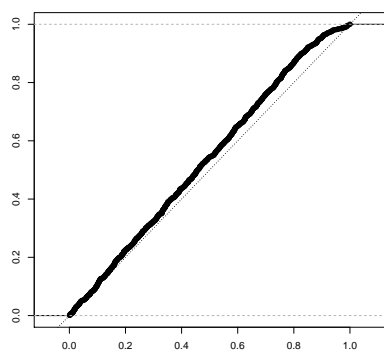
Figure E1: Distributions of p -values of REML-based tests
All plots are empirical CDFs of p -values, numbered as successive rows in Table 1.

Linear mixed model

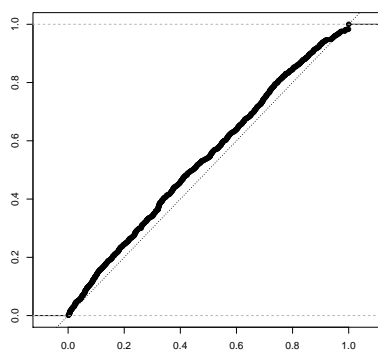


Poisson GLMM

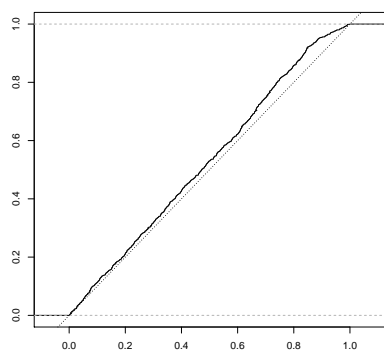
Case 1



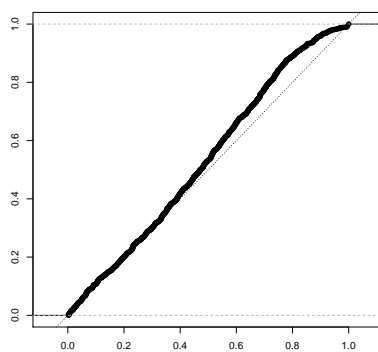
Case 2



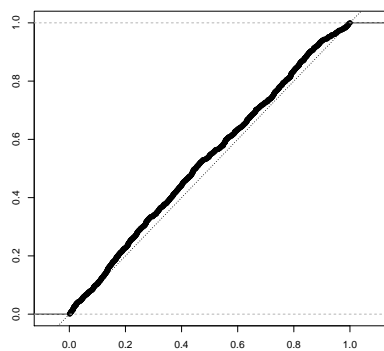
Case 3



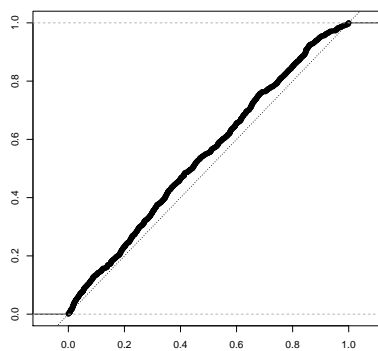
Case 4



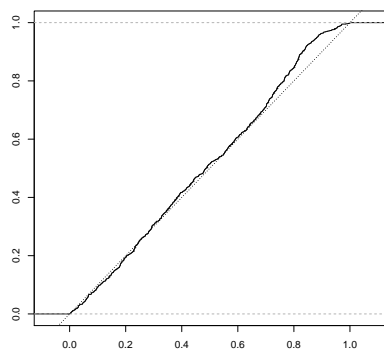
Case 5



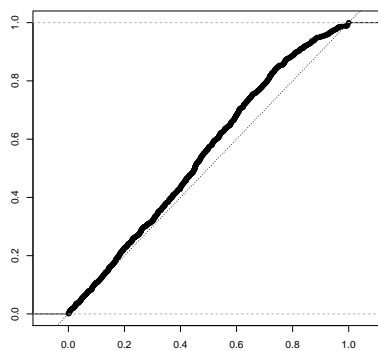
Case 6



Case 7

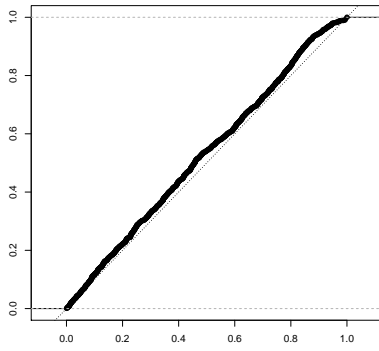


Case 8

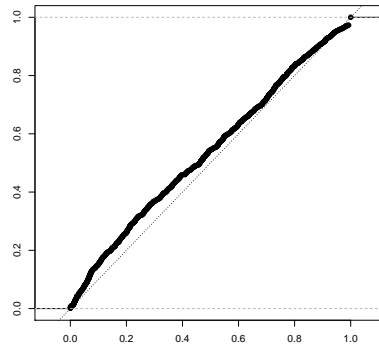


Binomial GLMM

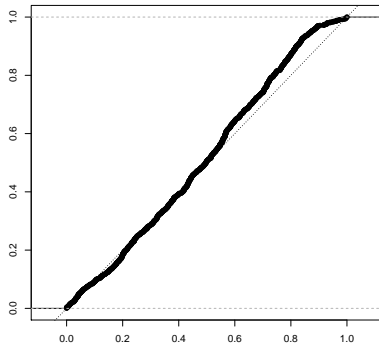
Case 1



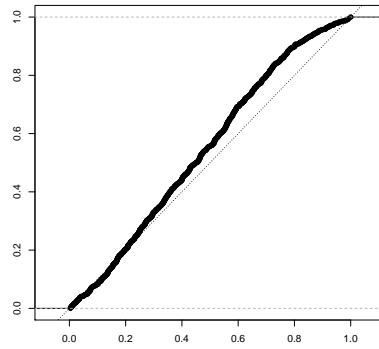
Case 2



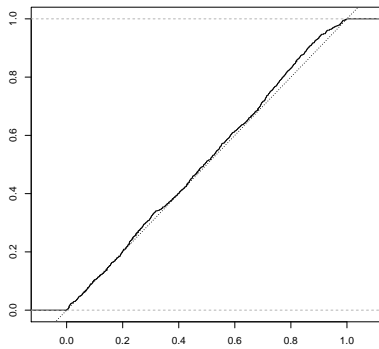
Case 3



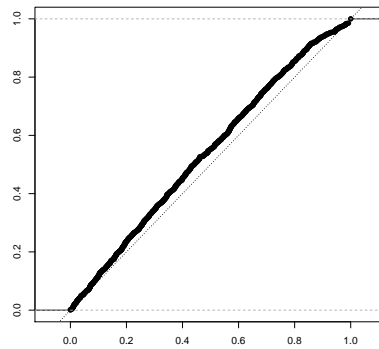
Case 4



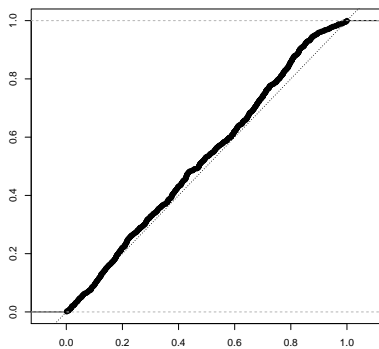
Case 5



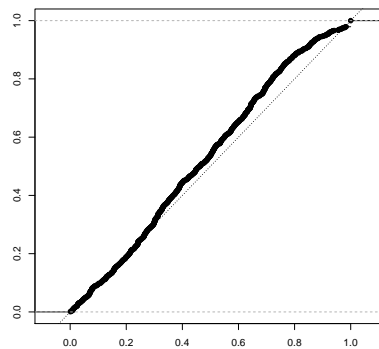
Case 6



Case 7



Case 8



Appendix F Binary GLMM

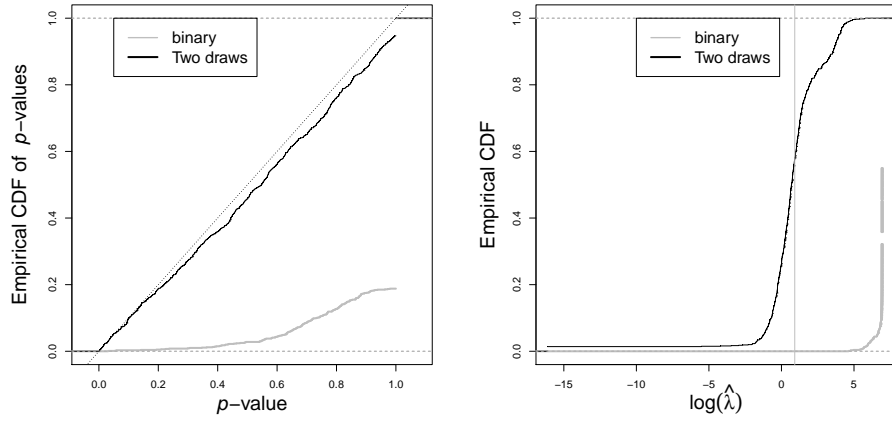


Figure F1: Comparison of ML analyses of a binary GLMM vs. a binomial GLMM with two draws in each location

Grey curves show the results of analyses (fit with the constraint $\hat{\lambda} < 1000$ for a simple graphical output) of binary data for parameters from the fourth row of the Tables, and black curves show the results for the same conditions, but with two draws in each spatial position. The true $\lambda = 2.5$ is marked by a vertical line.

Appendix G Experiments with available R procedures

This section describes experiments with an implementation of the Matérn model suitable for use with R procedures such as `lme`, `glmmPQL` or `gls`. In particular, the `glmmPQL` function from the R package `MASS` can fit GLMMs with spatial correlations, and provides estimates based on penalized-quasi likelihood approximations (PQL), which we compared to estimates obtained by the methods implements in `spaMM`. The details are fastidious but illustrate problems commonly encountered in practice with existing procedures. We do not discuss here procedures that would depend on additional user input, such as specification of the number of iterations of a Markov chain or of prior distributions.

Input and output

Syntactically, `glmmPQL` does not fit models where the only random effect is spatial. It shares this with the `lme` function from the `nlme` package, on which `glmmPQL` is based. It has been suggested to add a spurious random effect with a single level in order to fit the model. Then, the call is of the form

```
glmmPQL(<fixed effect formula>, random=~1|dummy,  
        correlation=corExp(form=~x+y|dummy), data=<data frame>)
```

where “dummy” is a spurious random effect with only one level for all the data analyzed. In a Gaussian LMM, the specification `corExp(form=~x+y|dummy)` means that there is an exponential correlation, with respect to the Euclidian distance between points with coordinates `x` and `y`, of the *residual* error among points within the same level of the conditioning dummy variable, that is, among all points in the present case. The “residual” variance reported by `lme` or `glmmPQL` is then the estimated variance of random effects, while the reported “intercept” variance is the variance of an uncorrelated residual error. In the Poisson or binomial GLMMs, the “residual” variance reported by `glmmPQL` is an overdispersion estimate for a quasipoisson or quasibinomial model.

The use of such a spurious random effect may look suspicious, but the fits it returns can be compared either to those obtained by analysing data that are duplicates of the original data, with two levels of the dummy variable, or to those obtained by our ML or PQL procedures. The `glmmPQL`-based procedure has also been described (Dormann et al., 2007, Appendix) as reproducing the results of the GLIMMIX procedure in SAS (which was also based on quasi-likelihood methods;

SAS, 2006). Further, for LMMs, `lme`, `glmmPQL` and the ML procedures should give consistent results as PQL estimation of fixed effects is equivalent to ML estimation in this case.

lme

Unexpectedly, `lme` systematically returned a practically null “intercept” variance in ML (as opposed to REML) fits. This `lme` result also differs from ML fits returned by our procedures, but the “residual” variance was correspondingly increased, so that sums of the two variances obtained by each ML implementation are similar. We then expect, and have checked, that this usage of `lme` gives results similar to the ML results on data simulated with low residual variance, namely the samples for the migration gene study design. We also implemented Matérn correlations within this framework. However, a substantial fraction of samples could then not be fitted due to lack of convergence of the optimization procedure called within `lme`.

If there are no repeated observations in given geographic locations, the residual variance can be modelled as a so-called nugget effect in the correlation function. However, we found that the optimization procedure almost always diverged when nugget estimation was attempted. Thus, one cannot analyze the simulated gaussian data sets from Table 1 under the model they were simulated from without additional programming. Moreover, this solution would not be appropriate for data sets with repeated observations in the same geographic location, because in this case the nugget and the residual variance are separable parameters.

glmmPQL

For GLMMs the PQL and ML results could be different, depending on the magnitude of random effects and the sample size. However, by comparing ML fits, `glmmPQL` fits and PQL/L fits, we found that `glmmPQL` performed much worse than PQL/L, which is very close to ML. For samples for which the `glmmPQL` fits converged, and for a low variance of random effects, estimates of fixed effects were close by both approaches and the `glmmPQL`-based analyses provided a good test of fixed effects. For a higher variance, `glmmPQL` fixed effect estimates depart from ML and PQL/L estimates, and the distribution of p -values was clearly distorted.

Estimation performance is more easily compared on larger samples, and we turned to samples simulated under the design of the onchocerciasis prevalence study for

such a comparison. In this case, all glmmPQL fits converged. ML and PQL/L fits are practically identical, with >0.9999 correlation of estimates of the different parameters, while PQL/L and glmmPQL differ as shown in Fig. G1. The performance of the tests is shown in the main text (Fig. 2, right). We checked that these discrepancies was not due to our implementation of the Matérn correlation for use with glmmPQL, by performing the same comparisons under a exponential correlation model (all other parameters unchanged), with similar results (Fig. G2).

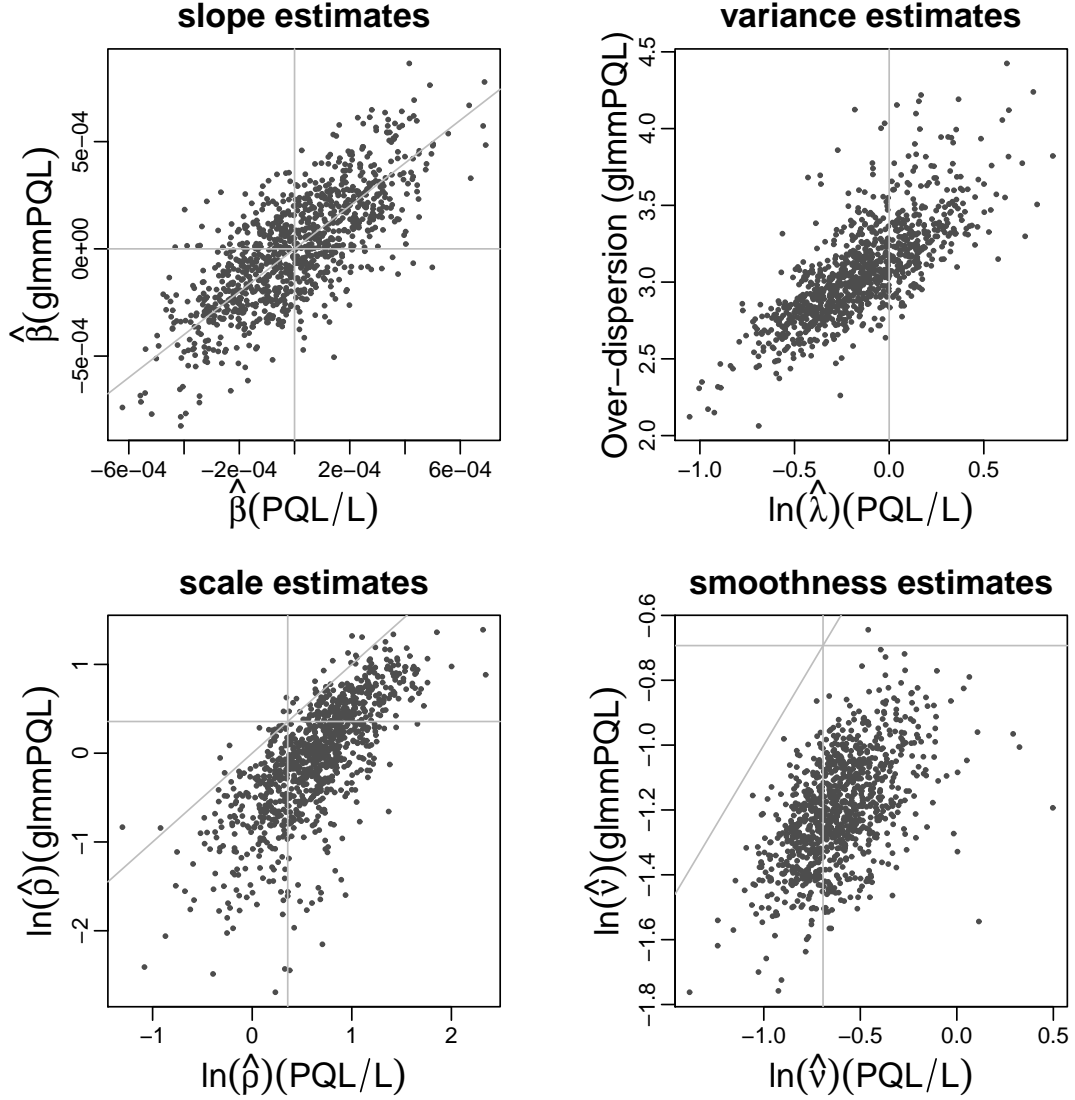


Figure G1: Joint distributions of parameter estimates by PQL/L and glmmPQL methods for data simulated according to the onchocerciasis study design. The diagonal line is the 1:1 line, and the horizontal and vertical gray lines mark the position of the true value, except for over-dispersion in the glmmPQL analysis, where true value is hard to defined in all but the simplest designs.

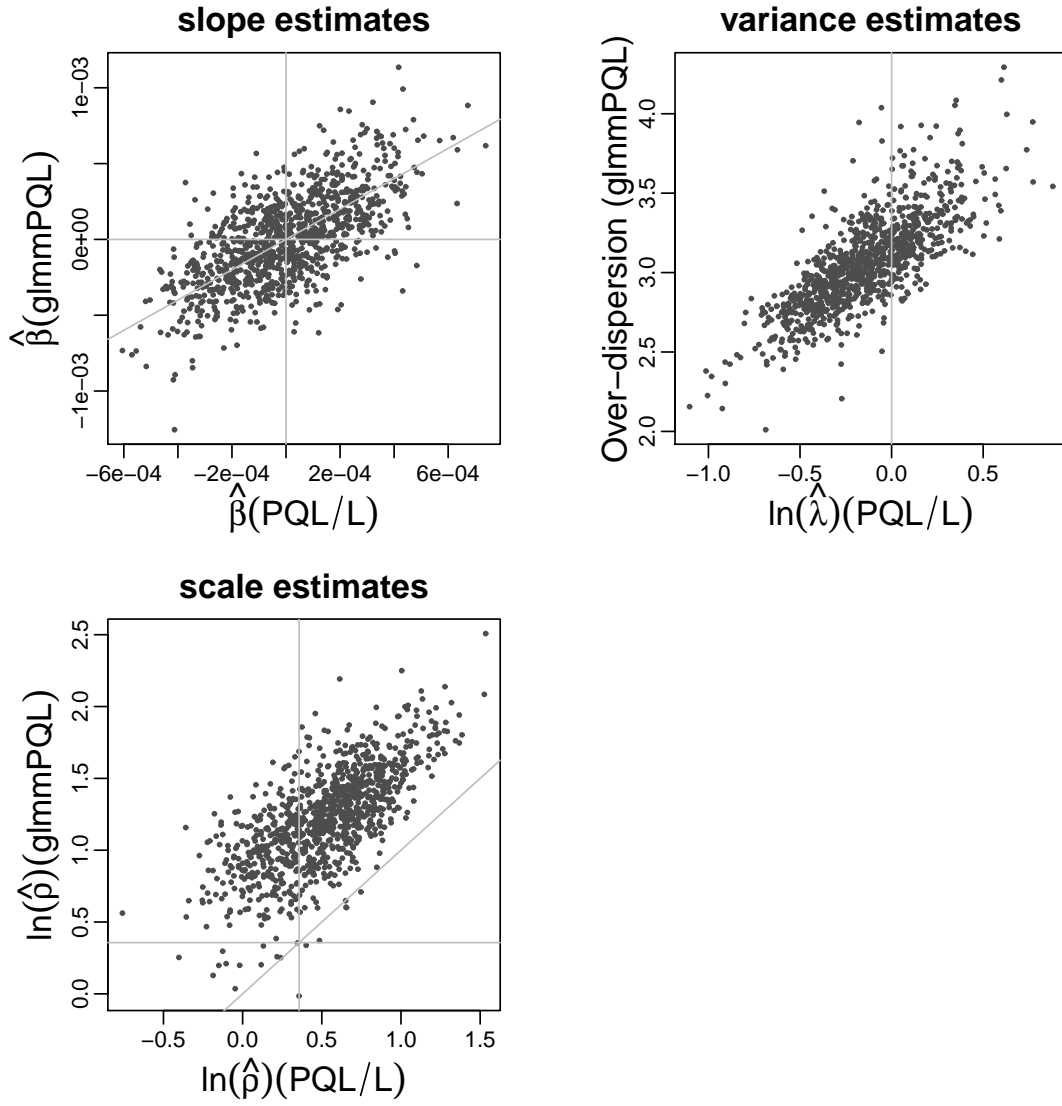


Figure G2: Same as Fig. G1 but for samples simulated and analyzed under an exponential correlation model

References

- Breslow, N. E., and D. G. Clayton. 1993. Approximate inference in generalized linear mixed models. *Journal of the American Statistical Association* **88**:9–25.
- Dormann, C. F., J. M. McPherson, M. B. Araújo, R. Bivand, J. Bolliger, G. Carl, R. G. Davies, A. Hirzel, W. Jetz, W. D. Kissling, I. Kühn, R. Ohlemüller, P. R. Peres-Neto, B. Reineking, B. Schröder, F. M. Schurr, and R. Wilson. 2007. Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. *Ecography* **30**:609–628.
- Lee, Y., and J. A. Nelder. 1996. Hierarchical generalized linear models. *Journal of the Royal Statistical Society B* **58**:619–678.
- Lee, Y., and J. A. Nelder. 2001. Hierarchical generalised linear models: A synthesis of generalised linear models, random-effect models and structured dispersions. *Biometrika* **88**:987–1006.
- Lee, Y., J. A. Nelder, and Y. Pawitan. 2006. *Generalized linear models with random effects: unified analysis via H-likelihood*. Chapman & Hall.
- Noh, M., and Y. Lee. 2007. REML estimation for binary data in GLMMs. *Journal of Multivariate Analysis* **98**:896–915.
- Pinheiro, J. C., and D. M. Bates. 2000. *Mixed-effects models in S and S-PLUS*. Springer Verlag, New York.
- SAS. 2006. The GLIMMIX procedure, June 2006. SAS Institute Inc, Cary, NC.
- Welham, S. J., and R. Thompson. 1997. Likelihood ratio tests for fixed model terms using residual maximum likelihood. *Journal of the Royal Statistical Society B* **59**:701–714.