Table Placeholders

Methods

Table 1: Overview of issues and recommendations for common classes of models. Correlation and distributions refer to predicted data from a fitted model, against which observed points are compared. A linear rotation refers to a multiplication of the simulated and observed data by a Cholesky decomposition of the estimated covariance matrix of the observed data, z'=Lz, as available in DHARMa.

Model class	Case studies	Issues and causes	Recommendation
Linear model	Linear model	No issues	Pearson residuals
Generalized linear model (GLM)	Skewed Gamma	Non-normality caused by response variable. Quantile residuals are needed if not approximately normal.	Quantile residual
Linear mixed model (LMM), Multivariate model	Random walk, Spatial LMM, Multinomial	Non-linear correlations caused by non-independence in observations.	Use a method that linearly decorrelates in order to scale to a unit iid normal. OSA Full Gaussian or simulation residuals with rotation.
Generalized linear mixed model (GLMM)	Spatial Poisson, Repeated measures Gamma	Non-normality and non-linear correlations caused by response variable and non-independence in observations.	Needs non-linear decorrelation and quantiles. OSA is only viable option.

Table 2: Linear Model Simulation: data generating models, parameter values, and mis-specifications.

Data Generating Model	Parameters	Mis-specified Model
$X_i \sim N(0,1)$ $\mu_{i,j} = X_i\beta$ $y_{i,j} \sim N(\mu_{i,j}, \sigma_y)$	$\beta = (4,-5)$ $\sigma_y = 1$	Data simulated with lognormal overdispersion: $\mu_{i,j} = X_i\beta + exp(\epsilon)$ $\epsilon \sim N(0,1)$ Data fit to model without drift term

Table 3: Mixed Model Simulation: data generating models, parameter values, and mis-specifications.

Data Generating Model	Parameters	Mis-specified Model
$X_i \sim N(0,1)$ $u_j \sim N(0,\sigma_u)$ $\mu_{i,j} = X_i\beta + u_j$ $y_{i,j} \sim N(\mu_{i,j},\sigma_y)$	$\beta = (4,-8)$ $\sigma_u = 2$ $\sigma_y = 0.5$	Data simulated with covariate term Data fit to model without covariate term
$X_i \sim Unif(-0.5,0.5)$ $u_j \sim N(0,\sigma_u)$ $\mu_{i,j} = X_i\beta + u_j$ $y_{i,j} \sim N(\mu_{i,j},\sigma_y)$	$\beta = (4,-8)$ $\sigma_u = 2$ $\sigma_y = 0.5$	Data simulated with covariate term Data fit to model without covariate term
$u_j \sim N(0,\sigma_u)$ $\mu_{i,j} = exp(\beta_0 + u_j)$ $y_{i,j} \sim Tweedie(\mu_{i,j},\phi,p)$	$\beta = 1.5$ $\sigma_u = 1.4$ $\phi = 1.4$ $p = 1.2$	Data simulated with random effect term Data fit to model without random effect term

Table 4: Randomwalk Simulation: data generating models, parameter values, and mis-specifications.

Data Generating Model	Parameters	Mis-specified Model
$\mu_i = u_{i-1} + a$	$a = 0.75$	
$u_i \sim N(\mu_i,\tau)$	$\tau = 1$	Data simulated with drift term, a
$y_i \sim N(u_i,\sigma)$	$\sigma = 1$	Data fit to model without drift term

Table 5: Spatial Simulation: data generating models, parameter values, and mis-specifications.

Data Generating Model	Parameters	Mis-specified Model
	spatial range = 50	
$\omega \sim GMRF(Q[\kappa,\sigma^2_\omega])$	$\kappa = \sqrt{8}/50$	
$\eta_i = \beta_0 + \omega_i$	$\sigma^2_\omega = 1$	Data simulated with $\exp(\omega_i)$
$y \sim N(\eta,\sigma_y)$	$\beta_0 = 1$	Data fit to model without covariate term
	$\sigma_y = 1$	
	spatial range = 50	
$\omega \sim GMRF(Q[\kappa,\sigma^2_\omega])$	$\kappa = \sqrt{8}/50$	
$\eta_i = \beta_0 + \omega_i$	$\sigma^2_\omega = 2$	Data simulated with random effect term
$y \sim Pois(exp(\eta))$	$\beta_0 = 0.5$	Data fit to model without random effect term

Results

Linear Model

Table 6: Linear Model. Type I error rates and Power evaluated for each analytical and simulation method for theoretical residuals using the KS normality test. Results are partitioned out by residual type (top to bottom).

	Overdispersion	
method	Type I Error	Power
Pearson	0.048	1
one-step Generic	0.048	1
one-step Gaussian	0.048	1
full Gaussian	0.048	1
cdf	0.048	1
Unconditional ecdf, Not Rotated	0.041	1
Conditional ecdf, Not Rotated	0.044	1

Table 7: Linear Model. Type I error rates and Power evaluated for each analytical and simulation method for estimated residuals using the KS normality test. Results are partitioned out by residual type (top to bottom).

	Overdispersion	
method	Type I Error	Power
Pearson	0	0.962
one-step Generic	0	0.963
one-step Gaussian	0	0.963
full Gaussian	0	0.963
cdf	0	0.963
Unconditional ecdf, Not Rotated	0	0.962
Conditional ecdf, Not Rotated	0	0.961

Mixed Model

Table 8: Mixed Model. Type I error rates and Power evaluated for each analytical and simulation method for theoretical residuals using the KS normality test. Results are partitioned out by model mis-specification (from left to right) and residual type (top to bottom).

	GLMM - Drop RE		LMM - Missing X Normal		LMM - Missing X Uniform	
method	Type I Error	Power	Type I Error	Power	Type I Error	Power
MCMC	0.042	1	0.053	1	0.051	1.000
Unconditional ecdf, Rotated	0.809	1	0.994	1	0.994	1.000
Unconditional ecdf, Not Rotated	0.991	1	0.999	1	0.999	0.999
Conditional ecdf, Rotated	0.995	1	0.996	1	0.996	1.000
Conditional ecdf, Not Rotated	0.035	1	0.047	1	0.047	1.000
Pearson	NA	NA	0.045	1	0.045	1.000
one-step Generic	NA	NA	0.043	1	0.043	1.000
one-step Gaussian	NA	NA	0.043	1	0.043	1.000
full Gaussian	NA	NA	0.043	1	0.043	1.000
cdf	NA	NA	0.043	1	0.043	1.000

Table 9: Mixed Model. Type I error rates and Power evaluated for each analytical and simulation method for estimated residuals using the KS normality test. Results are partitioned out by model mis-specification (from left to right) and residual type (top to bottom).

	GLMM - Drop RE		LMM - Missing X Normal		LMM - Missing X Uniform	
method	Type I Error	Power	Type I Error	Power	Type I Error	Power
MCMC	0.050	0.667	0.035	0.257	0.044	0.858
Unconditional ecdf, Rotated	0.935	0.999	0.998	0.999	0.998	1.000
Unconditional ecdf, Not Rotated	0.710	0.654	0.916	0.074	0.916	0.067
Conditional ecdf, Rotated	0.999	1.000	0.999	0.998	0.997	1.000
Conditional ecdf, Not Rotated	0.000	0.664	0.000	0.203	0.000	0.780
Pearson	NA	NA	0.000	0.210	0.000	0.780
one-step Generic	NA	NA	0.034	0.278	0.029	0.901
one-step Gaussian	NA	NA	0.034	0.278	0.029	0.901
full Gaussian	NA	NA	0.034	0.278	0.029	0.901
cdf	NA	NA	0.034	0.278	0.029	0.901

Randomwalk

Table 10: Randomwalk Model. Type I error rates and Power evaluated for each analytical and simulation method for theoretical residuals using the KS normality test. Results are partitioned out by model misspecification (from left to right) and residual type (top to bottom).

	mu0	
method	Type I Error	Power
Pearson	0.039	1.000
one-step Generic	0.038	1.000
one-step Gaussian	0.038	1.000
full Gaussian	0.038	1.000
cdf	0.041	1.000
MCMC	0.050	0.055
Unconditional ecdf, Rotated	0.058	1.000
Unconditional ecdf, Not Rotated	0.987	1.000
Conditional ecdf, Rotated	0.053	1.000
Conditional ecdf, Not Rotated	0.046	1.000

Table 11: Randomwalk Model. Type I error rates and Power evaluated for each analytical and simulation method for estimated residuals using the KS normality test. Results are partitioned out by model misspecification (from left to right) and residual type (top to bottom).

	mu0	
method	Type I Error	Power
Pearson	0.128	1.000
one-step Generic	0.000	1.000
one-step Gaussian	0.000	1.000
full Gaussian	0.000	1.000
cdf	0.000	1.000
MCMC	0.045	0.042
Unconditional ecdf, Rotated	0.000	1.000
Unconditional ecdf, Not Rotated	0.996	1.000
Conditional ecdf, Rotated	0.080	1.000
Conditional ecdf, Not Rotated	0.113	1.000

Spatial

Table 12: Spatial Model. Type I error rates and Power evaluated for each analytical and simulation method for theoretical residuals using the KS normality test. Results are partitioned out by model mis-specification (from left to right) and residual type (top to bottom).

	GLMM - Drop RE		LMM - Lognormal RE	
method	Type I Error	Power	Type I Error	Power
Pearson	0.452	1.000	0.046	1.000
one-step Generic	0.033	0.979	0.042	0.286
cdf	0.047	0.979	0.042	0.292
MCMC	0.039	0.979	0.042	0.088
Unconditional ecdf, Rotated	0.381	0.987	0.044	0.372
Unconditional ecdf, Not Rotated	0.662	0.973	0.689	1.000
Conditional ecdf, Rotated	0.081	0.989	0.049	1.000
Conditional ecdf, Not Rotated	0.037	0.976	0.039	1.000
one-step Gaussian	NA	NA	0.042	0.286
full Gaussian	NA	NA	0.042	0.286

Table 13: Spatial Model. Type I error rates and Power evaluated for each analytical and simulation method for estimated residuals using the KS normality test. Results are partitioned out by model mis-specification (from left to right) and residual type (top to bottom).

	GLMM - Drop RE		LMM - Lognormal RE	
method	Type I Error	Power	Type I Error	Power
Pearson	0.859	1.000	0.183	0.655
one-step Generic	0.007	0.965	0.013	0.366
cdf	0.006	0.964	0.015	0.368
MCMC	0.039	0.969	0.043	0.297
Unconditional ecdf, Rotated	0.193	0.970	0.027	0.325
Unconditional ecdf, Not Rotated	0.270	0.966	0.272	0.636
Conditional ecdf, Rotated	0.345	0.969	0.109	0.575
Conditional ecdf, Not Rotated	0.415	0.960	0.186	0.660
one-step Gaussian	NA	NA	0.013	0.366
full Gaussian	NA	NA	0.013	0.366