persistencePits v7.0

Matthew E. Verbyla

6/16/2022

## Load packages

First, load the required packages.

library(ggplot2)  
library(gridExtra)  
library("PerformanceAnalytics")  
library("Hmisc")  
library("knitr")  
library("bbmle")

## Load and explore data

Now, we will load and explore the persistence data, which are stored on the GWPP K2P Data Portal under the following dataset: <http://data.waterpathogens.org/dataset/persistence-treatment>.

The next chunk of code explores the data set to quantify the total number of data points, experiments, and the breakdown of lab-based studies, field-based studies, and the number of experiments done with different types of pathogen groups.

nrow(persist) # total number of data points

## [1] 1382

length(unique(persist$experiment\_id)) # total number of experiments

## [1] 243

length(persist[persist$location\_comment=="Laboratory-based study",]$experiment\_id) # number of data points from lab-based studies

## [1] 1078

length(unique(persist[persist$location\_comment=="Laboratory-based study",]$experiment\_id)) # number of lab-based experiments

## [1] 190

length(persist[persist$location\_comment!="Laboratory-based study",]$experiment\_id) # number of data points from field-based studies

## [1] 304

length(unique(persist[persist$location\_comment!="Laboratory-based study",]$experiment\_id)) # number of field-based experiments

## [1] 53

length(unique(persist[persist$microbial\_group=="Virus",]$experiment\_id)) # number of experiments with viruses

## [1] 32

length(unique(persist[persist$microbial\_group=="Bacteria",]$experiment\_id)) # number of experiments with bacteria

## [1] 147

length(unique(persist[persist$microbial\_group=="Protozoa",]$experiment\_id)) # number of experiments with protozoa

## [1] 8

length(unique(persist[persist$microbial\_group=="Helminth",]$experiment\_id)) # number of experiments with helminths

## [1] 56

unique(persist[persist$microbial\_group=="Virus",]$microorganism\_common\_name) # types of viruses studied

## [1] "Bacteriophage φX" "Bacteriophage MS2"   
## [3] "Salmonella typhimurium phages" "MS2 coliphage"   
## [5] "PhiX174 coliphage" "T4 coliphage"   
## [7] "Adenovirus" "Salmonella typhimurium phage 28B"  
## [9] "Rotavirus"

unique(persist[persist$microbial\_group=="Bacteria",]$microorganism\_common\_name) # types of bacteria studied

## [1] "Fecal Coliforms" "Enterococcus faecalis"   
## [3] "Salmonella enterica Typhimurium" "Escherichia Coli"   
## [5] "Fecal coliforms" "Streptococcus"   
## [7] "Salmonella" "Total coliform"   
## [9] "Eschericia Coli" "Sulfur reducing clostridia"   
## [11] "Escherichia coli" "Enterococcus"   
## [13] "Salmonella spp" "Thermotolerant coliform"   
## [15] "Enterococci" "Salmonella senftenberg"   
## [17] "Total coliforms" "Salmonella spp."   
## [19] "Shigella sonnei" "Shigella flexneri"   
## [21] "Escherichia coli "

unique(persist[persist$microbial\_group=="Protozoa",]$microorganism\_common\_name) # types of protozoa studied

## [1] "Entamoeba cysts" "Cryptosporidium" "Giardia"

unique(persist[persist$microbial\_group=="Helminth",]$microorganism\_common\_name) # types of helminths studied

## [1] "Ascaris" "Ascaris suum" "Ascaris eggs"   
## [4] "Ascaris ova" "Ascaris lumbricoides"

## Calculate the decay rate coefficients

The next chunk of code sets up some variables that will be used in an upcoming "FOR" loop, which will be used to calculate the decay rate coefficients and other statistics associated with the fit of the log-linear and JM2 models.

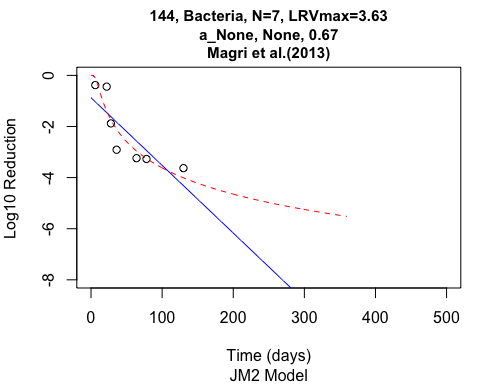
N<-length(unique(persist$experiment\_id)) # the total number of experiments  
M<-length(persist$experiment\_id) # the total number of data points  
persist$ind<-NA  
for(i in 1:M){  
 # this creates a unique index variable to distinguish each independent experiment  
 persist$ind[i]<-which(data.frame(unique(persist$experiment\_id))==persist$experiment\_id[i])  
}  
# here we create lots of NULL variables, which later we will populate values for in the upcoming "for" loop  
k<-NULL;k1<-NULL;k2<-NULL;intrcpt<-NULL;group<-NULL;addit<-NULL;urine<-NULL;urea<-NULL;moist<-NULL;temp<-NULL;pH<-NULL;r2<-NULL;t99<-NULL;t99jm<-NULL;authors<-NULL;num<-NULL;aicLL<-NULL;aicJM<-NULL;experiment<-NULL;maxLRV<-NULL;maxTime<-NULL;mcn<-NULL;best<-NULL;sterr<-NULL;sigme<-NULL  
par(mfrow=c(3,6)) #this creates the panel layout for the plots used for creating Figure S1

for(i in 1:N){ #in this loop, we calculate the decay rate coefficients for the log-linear and JM2 models  
 time<-persist[persist$ind==i,]$time\_days # get the time for the present experiment  
 lnrv<-persist[persist$ind==i,]$ln\_removal # get the ln reduction for the present experiment  
 log10\_reduction<-persist[persist$ind==i,]$log10\_reduction # get the log10 reduction for the present experiment  
 # since we calculated the ln reduction, then equation gets algebraically rearranged like this:  
 # Ct = Co\*exp(-k\*t)   
 # ln(Ct/Co) = -k\*t  
   
 # first, we fit the model like this to express the decay rate coefficient, k, on a LN scale  
 fit<-lm(lnrv~time) # lnrv = ln(Ct/Co), so our linear model is lnrv~time  
 # then, we fit the model like this to create plots and calculate T99 values more easily later on  
 fit2<-lm(log10\_reduction~time)  
   
 # next, we set up the JM2 model  
 JM2<-function(k1,k2,sigma){  
 R = lnrv - log(1/(1+exp(k1+k2\*log(time)))) # this is the equation for the JM2 model  
 R = suppressWarnings(dnorm(R,0,sigma,log=T)) # this is the standard deviation of the normally distributed difference between measured and modeled LRVs  
 -sum(R) # this is the negative log likelihood, which must be maximized (using the MLE method)  
 }  
   
 # in order to use MLE to find the best values for the JM2 model, we have to provide some halfway decent guesses for k1, k2, and sigma  
 # through trial and error, we found that starting guesses of k1=-5, k2=1, and sigma=3 generally produced converging results for most experiments; the following are exceptions  
 if(fit$coefficients[2]<=-10){k1guess=-fit$coefficients[2];k2guess=-fit$coefficients[2];siguess=3}else{  
 if(fit$coefficients[2]>=-0.0001){k1guess=-5;k2guess=0.01;siguess=2}else{  
 if(i==48|i==181|i==186){k1guess=-0.1;k2guess=1;siguess=1}else{  
 if(i==115){k1guess=4;k2guess=100;siguess=1}else{  
 if(i==112|i==113|i==105|i==114){k1guess=5;k2guess=1;siguess=1}else{  
 if(i==180){k1guess=1;k2guess=10;siguess=1}else{  
 k1guess=-5;k2guess=1;siguess=3  
 }  
 }  
 }  
 }  
 }  
 }  
  
 if(length(time)>3){ # it only calculates the k1, k2, and sigma values for the JM2 model if there are more than 3 time points in the experiment  
 fit.jm<-mle2(minuslogl=JM2,start=list(k1=k1guess,k2=k2guess,sigma=siguess),optimizer="nlminb") # uses unbounded optimization using portable Fortran programs for numerical computation (PORT) routines   
 # Reference: Fox et al. (1978). “The PORT Mathematical Subroutine Library.” ACM Transactions on Mathematical Software, 4(2), 104–126. doi:10.1145/355780.355783  
 }else{fit.jm<-NA}  
   
 ## NEXT WE STORE THE VALUES OF THE VARIABLES AND STATISTICS WE WANT TO STORE FOR EACH EXPERIMENT  
 # here, we store the basic information about the experiment (e.g., the design, operational, and environmental parameters)   
 authors[i]<-as.character(unique(persist[persist$ind==i,]$authors)) # the names of the authors of the paper where the experiment's data were published  
 experiment[i]<-as.numeric(unique(persist[persist$ind==i,]$experiment\_id)) # the unique experiment ID  
 num[i]<-length(time) # number of time points in the experiment  
 maxTime[i]<-max(time) # the maximum time point for which data were collected in the experiment  
 maxLRV[i]<-as.numeric(max(persist[persist$ind==i,]$log10\_reduction)) # the maximum log10 reduction value observed in the experiment  
 group[i]<-as.character(unique(persist[persist$ind==i,]$microbial\_group)) # the microbial group used in the experiment (viruses, bacteria, protozoa, helminths)  
 mcn[i]<-as.character(persist[persist$ind==i,]$microorganism\_common\_name) # a more specific description of the microbial group  
 addit[i]<-as.character(unique(persist[persist$ind==i,]$additive)) # the type of additive applied to the fecal sludge  
 urine[i]<-as.character(unique(persist[persist$ind==i,]$urine)) # whether the experiment was done with feces only or feces mixed with urine (excreta)  
 moist[i]<-as.numeric(median(persist[persist$ind==i,]$moisture\_content\_percent)) # the average percent moisture content of the fecal sludge during the experiment  
 temp[i]<-as.numeric(median(persist[persist$ind==i,]$temperature\_celsius)) # the average temperature of the fecal sludge during the experiment  
 pH[i]<-as.numeric(median(persist[persist$ind==i,]$pH)) # the average pH of the fecal sludge during the experiment  
 urea[i]<-as.character(unique(persist[persist$ind==i,]$urea)) # whether or not urea (or stored urine) was added to the fecal sludge during the experiment  
  
 # here, we store the statistics for the fit to the log-linear model  
 k[i]<-fit$coefficients[2] # the log-linear decay rate coefficient  
 intrcpt[i]<-fit$coefficients[1] # the log-linear model Y-intercept  
 out<-summary(fit); sterr[i]<-out$coefficients[1,2] # this is the standard error of the decay rate coefficient  
 t99[i]<-(2-fit2$coefficients[1])/fit2$coefficients[2] # the predicted T99 value for the log-linear model  
 r2[i]<-summary(fit)$r.squared # coefficient of determination for the log-linear model fit  
 aicLL[i]<-AIC(fit) # the AIC value for the log-linear model fit  
   
 # and here, we store the statistics for the fit to the JM2 model  
 if(!is.na(fit.jm)){k1[i]<-coef(fit.jm)["k1"]} # the 1st JM2 model decay rate coefficient  
 if(!is.na(fit.jm)){k2[i]<-coef(fit.jm)["k2"]} # the 2nd JM2 model decay rate coefficient  
 if(!is.na(fit.jm)){sigme[i]<-coef(fit.jm)["sigma"]} # the sigma value for the fit of the JM2 model (measure of uncertainty)  
 t99jm[i]<-exp((log(99)-k1[i])/k2[i]) # the predicted T99 value for the JM2 model  
 if(length(time)>3){aicJM[i]<-AIC(fit.jm)}else{aicJM[i]<-NA} # the AIC value for the JM2 model fit  
   
 # this stores text describing which of the two models had the better fit (Chick Log-Linear Model or JM2 Model), based on the AIC values calculated  
 if(is.na(aicJM[i])){best[i]<-"Chick Model"}else{if(aicLL[i]<aicJM[i]){best[i]<-"Chick Model"}else{best[i]<-"JM2 Model"}}  
  
 # this next chunk of script creates the plot for each panel of Figure S1 (new plot created for each experiment during each round of the FOR loop)  
 par(mar=c(5,4,3.5,1)) # sets the margins so that the plot looks nice  
 plot(time,-log10\_reduction,ylim=c(-8,0),xlim=c(0,500),ylab="Log10 Reduction",xlab="Time (days)",main=paste(experiment[i],", ",group[i],", N=",num[i],", LRVmax=",round(maxLRV[i],2),"\n",addit[i],", ",urea[i],", ",round(r2[i],2),"\n",authors[i],sep=""),sub=best[i],cex.main=0.95)  
 lines(seq(0,360),-fit2$coefficients[2]\*seq(0,360)-fit2$coefficients[1],col="blue")  
 if(length(time)>3){lines(seq(0,360),log10(1/(1+exp(coef(fit.jm)["k1"]+coef(fit.jm)["k2"]\*log(seq(0,360))))),lty=2,col="red")}  
}

# now we will create a data frame with all the values from all experiments  
kPit<-data.frame(experiment=experiment,authors=authors,microbial\_group=group,mcn=mcn,k=k,se.k=sterr,k1=k1,k2=k2,se.jm=sigme,int=intrcpt,aicLL=aicLL,aicJM=aicJM,num=num,additive=addit,urea=urea,urine=urine,moisture=moist,temp=temp,pH=pH,r2=r2,maxLRV=maxLRV,t99=t99,t99jm=t99jm,bestModel=best)  
kPit$k<-round(kPit$k,4)  
kPit$k1<-round(kPit$k1,2)  
kPit$k2<-round(kPit$k2,2)  
kPit.demo<-kPit[c(1,17,126,236),c("experiment","authors","mcn","k","k1","k2")]  
kPit.demo

## experiment authors mcn k  
## 1 128 Endale et al.(2012) Ascaris -0.0656  
## 17 144 Magri et al.(2013) Salmonella enterica Typhimurium -0.0610  
## 126 253 Decrey and Kohn (2017) MS2 coliphage -1.0044  
## 236 364 Graham et al. (2003) Cryptosporidium -0.0149  
## k1 k2  
## 1 -6.79 2.48  
## 17 -7.38 3.41  
## 126 -7.19 7.99  
## 236 NA NA

The previous chunk of code hid the plots that were created for Figure S1. Here is a demonstrative plot for experiment 144.



Here is a summary of which model provided the better fit for each experiment

kPit[,c("experiment","authors","mcn","bestModel")]

## experiment authors mcn  
## 1 128 Endale et al.(2012) Ascaris  
## 2 129 Endale et al.(2012) Ascaris  
## 3 130 Endale et al.(2012) Ascaris  
## 4 131 Endale et al.(2012) Ascaris  
## 5 132 Endale et al.(2012) Fecal Coliforms  
## 6 133 Endale et al.(2012) Fecal Coliforms  
## 7 134 Endale et al.(2012) Fecal Coliforms  
## 8 135 Endale et al.(2012) Fecal Coliforms  
## 9 136 Magri et al.(2013) Enterococcus faecalis  
## 10 137 Magri et al.(2013) Enterococcus faecalis  
## 11 138 Magri et al.(2013) Enterococcus faecalis  
## 12 139 Magri et al.(2013) Enterococcus faecalis  
## 13 140 Magri et al.(2013) Enterococcus faecalis  
## 14 141 Magri et al.(2013) Salmonella enterica Typhimurium  
## 15 142 Magri et al.(2013) Salmonella enterica Typhimurium  
## 16 143 Magri et al.(2013) Salmonella enterica Typhimurium  
## 17 144 Magri et al.(2013) Salmonella enterica Typhimurium  
## 18 145 Magri et al.(2013) Salmonella enterica Typhimurium  
## 19 146 Magri et al.(2013) Bacteriophage φX  
## 20 147 Magri et al.(2013) Bacteriophage φX  
## 21 148 Magri et al.(2013) Bacteriophage φX  
## 22 149 Magri et al.(2013) Bacteriophage φX  
## 23 150 Magri et al.(2013) Bacteriophage φX  
## 24 151 Magri et al.(2013) Bacteriophage φX  
## 25 152 Magri et al.(2013) Bacteriophage MS2  
## 26 153 Magri et al.(2013) Bacteriophage MS2  
## 27 154 Magri et al.(2013) Bacteriophage MS2  
## 28 155 Magri et al.(2013) Bacteriophage MS2  
## 29 156 Magri et al.(2013) Bacteriophage MS2  
## 30 157 Magri et al.(2013) Bacteriophage MS2  
## 31 158 Chien et al. (2002) Salmonella typhimurium phages  
## 32 159 Chien et al. (2002) Salmonella typhimurium phages  
## 33 160 Chien et al. (2002) Salmonella typhimurium phages  
## 34 161 Chien et al. (2002) Salmonella typhimurium phages  
## 35 162 Chien et al. (2002) Ascaris suum  
## 36 163 Chien et al. (2002) Ascaris suum  
## 37 164 Chien et al. (2002) Ascaris suum  
## 38 165 Fidjeland et al. (2013) Ascaris eggs  
## 39 166 Fidjeland et al. (2013) Ascaris eggs  
## 40 167 Fidjeland et al. (2013) Ascaris eggs  
## 41 168 Fidjeland et al. (2013) Ascaris eggs  
## 42 169 Fidjeland et al. (2013) Ascaris eggs  
## 43 170 Fidjeland et al. (2013) Ascaris eggs  
## 44 171 Fidjeland et al. (2013) Ascaris eggs  
## 45 172 Jensen et al. (2009) Ascaris eggs  
## 46 173 Ogunyoku et al. (2016) Escherichia Coli  
## 47 174 Ogunyoku et al. (2016) Escherichia Coli  
## 48 175 Ogunyoku et al. (2016) Escherichia Coli  
## 49 176 Ogunyoku et al. (2016) MS2 coliphage  
## 50 177 Ogunyoku et al. (2016) MS2 coliphage  
## 51 178 Ogunyoku et al. (2016) MS2 coliphage  
## 52 179 Ogunyoku et al. (2016) Ascaris  
## 53 180 Ogunyoku et al. (2016) Ascaris  
## 54 181 Ogunyoku et al. (2016) Ascaris  
## 55 182 Nordin et al. (2009) Ascaris  
## 56 183 Nordin et al. (2009) Ascaris  
## 57 184 Nordin et al. (2009) Ascaris  
## 58 185 Nordin et al. (2009) Ascaris  
## 59 186 Nordin et al. (2009) Ascaris  
## 60 187 Nordin et al. (2009) Ascaris  
## 61 188 Nordin et al. (2009) Ascaris  
## 62 189 Nordin et al. (2009) Ascaris  
## 63 190 McKinley et al. (2012) Ascaris ova  
## 64 191 McKinley et al. (2012) Ascaris ova  
## 65 192 McKinley et al. (2012) Ascaris ova  
## 66 193 McKinley et al. (2012) Ascaris ova  
## 67 194 McKinley et al. (2012) Ascaris ova  
## 68 195 McKinley et al. (2012) Ascaris ova  
## 69 196 Pompeo et al. (2016) Ascaris eggs  
## 70 197 Pompeo et al. (2016) Ascaris eggs  
## 71 198 Pompeo et al. (2016) Ascaris eggs  
## 72 199 Pompeo et al. (2016) Ascaris eggs  
## 73 200 Pompeo et al. (2016) Ascaris eggs  
## 74 201 Pompeo et al. (2016) Ascaris eggs  
## 75 202 Yin et al. (2016) Fecal coliforms  
## 76 203 Yin et al. (2016) Streptococcus  
## 77 204 Yin et al. (2016) Salmonella  
## 78 205 Yin et al. (2016) Fecal coliforms  
## 79 206 Yin et al. (2016) Streptococcus  
## 80 207 Yin et al. (2016) Salmonella  
## 81 208 Yin et al. (2016) Fecal coliforms  
## 82 209 Yin et al. (2016) Streptococcus  
## 83 210 Yin et al. (2016) Salmonella  
## 84 211 Yin et al. (2016) Fecal coliforms  
## 85 212 Yin et al. (2016) Streptococcus  
## 86 213 Yin et al. (2016) Salmonella  
## 87 214 Yin et al. (2016) Fecal coliforms  
## 88 215 Yin et al. (2016) Streptococcus  
## 89 216 Yin et al. (2016) Salmonella  
## 90 217 Yin et al. (2016) Fecal coliforms  
## 91 218 Yin et al. (2016) Streptococcus  
## 92 219 Yin et al. (2016) Salmonella  
## 93 220 Yin et al. (2016) Fecal coliforms  
## 94 221 Yin et al. (2016) Streptococcus  
## 95 222 Yin et al. (2016) Salmonella  
## 96 223 Sossou et al. (2014) Total coliform  
## 97 224 Sossou et al. (2014) Eschericia Coli  
## 98 225 Sossou et al. (2014) Streptococcus  
## 99 226 Sossou et al. (2014) Sulfur reducing clostridia  
## 100 227 Sossou et al. (2014) Entamoeba cysts  
## 101 228 Sossou et al. (2014) Ascaris eggs  
## 102 229 Darimani et al.(2015) Escherichia coli  
## 103 230 Darimani et al.(2015) Escherichia coli  
## 104 231 Darimani et al.(2015) Escherichia coli  
## 105 232 Darimani et al.(2015) Escherichia coli  
## 106 233 Darimani et al.(2015) Escherichia coli  
## 107 234 Darimani et al.(2015) Escherichia coli  
## 108 235 Darimani et al.(2015) Escherichia coli  
## 109 236 Darimani et al.(2015) Escherichia coli  
## 110 237 Darimani et al.(2015) Escherichia coli  
## 111 238 Darimani et al.(2015) Enterococcus  
## 112 239 Darimani et al.(2015) Enterococcus  
## 113 240 Darimani et al.(2015) Enterococcus  
## 114 241 Darimani et al.(2015) Enterococcus  
## 115 242 Darimani et al.(2015) Enterococcus  
## 116 243 Darimani et al.(2015) Enterococcus  
## 117 244 Darimani et al.(2015) Enterococcus  
## 118 245 Darimani et al.(2015) Enterococcus  
## 119 246 Darimani et al.(2015) Enterococcus  
## 120 247 Darimani et al.(2015) Enterococcus  
## 121 248 Darimani et al.(2015) Enterococcus  
## 122 249 Darimani et al.(2015) Enterococcus  
## 123 250 Darimani et al.(2015) Escherichia coli  
## 124 251 Darimani et al.(2015) Escherichia coli  
## 125 252 Darimani et al.(2015) Escherichia coli  
## 126 253 Decrey and Kohn (2017) MS2 coliphage  
## 127 254 Decrey and Kohn (2017) T4 coliphage  
## 128 255 Decrey and Kohn (2017) Adenovirus  
## 129 256 Fidjeland et al. (2016) Ascaris  
## 130 257 Fidjeland et al. (2016) Ascaris  
## 131 258 Fidjeland et al. (2016) Ascaris  
## 132 259 Fidjeland et al. (2016) Ascaris  
## 133 260 Fidjeland et al. (2016) Ascaris  
## 134 261 Fidjeland et al. (2016) Ascaris  
## 135 262 Fidjeland et al. (2016) Ascaris  
## 136 263 Fidjeland et al. (2016) Ascaris  
## 137 264 Fidjeland et al. (2016) Ascaris  
## 138 265 Fidjeland et al. (2016) Ascaris  
## 139 266 Fidjeland et al. (2016) Ascaris  
## 140 267 Fidjeland et al. (2016) Ascaris  
## 141 268 Fidjeland et al. (2016) Salmonella spp  
## 142 269 Fidjeland et al. (2016) Salmonella spp  
## 143 270 Fidjeland et al. (2016) Salmonella spp  
## 144 271 Fidjeland et al. (2016) Salmonella spp  
## 145 272 Fidjeland et al. (2016) Salmonella spp  
## 146 273 Fidjeland et al. (2016) Salmonella spp  
## 147 274 Fidjeland et al. (2016) Salmonella spp  
## 148 275 Fidjeland et al. (2016) Salmonella spp  
## 149 276 Fidjeland et al. (2016) Salmonella spp  
## 150 277 Fidjeland et al. (2016) Salmonella spp  
## 151 278 Fidjeland et al. (2016) Salmonella spp  
## 152 279 Fidjeland et al. (2016) Salmonella spp  
## 153 280 Tonner-Klank et al. (2007) Salmonella typhimurium phage 28B  
## 154 281 Tonner-Klank et al. (2007) Thermotolerant coliform  
## 155 282 Tonner-Klank et al. (2007) Enterococci  
## 156 283 Tonner-Klank et al. (2007) Salmonella senftenberg  
## 157 284 Tonner-Klank et al. (2007) Salmonella typhimurium phage 28B  
## 158 285 Tonner-Klank et al. (2007) Salmonella senftenberg  
## 159 286 Tonner-Klank et al. (2007) Thermotolerant coliform  
## 160 287 Tonner-Klank et al. (2007) Enterococci  
## 161 288 Odey et al. (2018) Fecal coliforms  
## 162 289 Odey et al. (2018) Fecal coliforms  
## 163 290 Odey et al. (2018) Fecal coliforms  
## 164 291 Odey et al. (2018) Fecal coliforms  
## 165 292 Niwagaba et al.(2009) Escherichia coli  
## 166 293 Niwagaba et al.(2009) Enterococci  
## 167 294 Niwagaba et al.(2009) Enterococci  
## 168 295 Niwagaba et al.(2009) Escherichia coli  
## 169 296 Niwagaba et al.(2009) Enterococci  
## 170 297 Niwagaba et al.(2009) Escherichia coli  
## 171 298 Niwagaba et al.(2009) Enterococci  
## 172 299 Niwagaba et al.(2009) Escherichia coli  
## 173 300 Anderson et al. (2015) Escherichia coli  
## 174 302 Anderson et al. (2015) Escherichia coli  
## 175 303 Anderson et al. (2015) Escherichia coli  
## 176 304 Anderson et al. (2015) Fecal coliforms  
## 177 305 Anderson et al. (2015) Fecal coliforms  
## 178 306 Anderson et al. (2015) Fecal coliforms  
## 179 307 Anderson et al. (2015) Fecal coliforms  
## 180 308 Anderson et al. (2015) Escherichia coli  
## 181 309 Anderson et al. (2015) Escherichia coli  
## 182 310 Anderson et al. (2015) Escherichia coli  
## 183 311 Anderson et al. (2015) Escherichia coli  
## 184 312 Anderson et al. (2015) Escherichia coli  
## 185 313 Anderson et al. (2015) Escherichia coli  
## 186 314 Anderson et al. (2015) Total coliforms  
## 187 315 Anderson et al. (2015) Total coliforms  
## 188 316 Anderson et al. (2015) Total coliforms  
## 189 317 Anderson et al. (2015) Total coliforms  
## 190 318 Anderson et al. (2015) Total coliforms  
## 191 319 Anderson et al. (2015) Total coliforms  
## 192 320 Nordin et al. (2009b) Salmonella spp.  
## 193 321 Nordin et al. (2009b) Salmonella spp.  
## 194 322 Nordin et al. (2009b) Salmonella spp.  
## 195 323 Hashemi et al. (2019) Escherichia coli  
## 196 324 Hashemi et al. (2019) Escherichia coli  
## 197 325 Hashemi et al. (2019) Escherichia coli  
## 198 326 Nakamura and Taylor (1965) Shigella sonnei  
## 199 327 Nakamura and Taylor (1965) Shigella sonnei  
## 200 328 Nakamura and Taylor (1965) Shigella flexneri  
## 201 329 Nakamura and Taylor (1965) Shigella sonnei  
## 202 330 Nakamura and Taylor (1965) Shigella sonnei  
## 203 331 Nakamura and Taylor (1965) Shigella flexneri  
## 204 332 Moe and Shirley (1982) Rotavirus  
## 205 333 Moe and Shirley (1982) Rotavirus  
## 206 334 Moe and Shirley (1982) Rotavirus  
## 207 335 Moe and Shirley (1982) Rotavirus  
## 208 336 Moe and Shirley (1982) Rotavirus  
## 209 337 Moe and Shirley (1982) Rotavirus  
## 210 338 Moe and Shirley (1982) Rotavirus  
## 211 339 Moe and Shirley (1982) Rotavirus  
## 212 340 Sossou et al.(2016) Escherichia coli   
## 213 341 Sossou et al.(2016) Escherichia coli   
## 214 342 Sossou et al.(2016) Escherichia coli   
## 215 343 Sossou et al.(2016) Escherichia coli   
## 216 344 Sossou et al.(2016) Escherichia coli   
## 217 345 Sossou et al.(2016) Enterococcus faecalis  
## 218 346 Sossou et al.(2016) Enterococcus faecalis  
## 219 347 Sossou et al.(2016) Enterococcus faecalis  
## 220 348 Sossou et al.(2016) Enterococcus faecalis  
## 221 349 Sossou et al.(2016) Enterococcus faecalis  
## 222 350 Sossou et al.(2016) Ascaris lumbricoides  
## 223 351 Sossou et al.(2016) Ascaris lumbricoides  
## 224 352 Sossou et al.(2016) Ascaris lumbricoides  
## 225 353 Sossou et al.(2016) Ascaris lumbricoides  
## 226 354 Sossou et al.(2016) Ascaris lumbricoides  
## 227 355 Berendes et al.(2015) Escherichia coli   
## 228 356 Berendes et al.(2015) Escherichia coli   
## 229 357 Berendes et al.(2015) Escherichia coli   
## 230 358 Berendes et al.(2015) Escherichia coli   
## 231 359 Berendes et al.(2015) Escherichia coli   
## 232 360 Berendes et al.(2015) Escherichia coli   
## 233 361 Berendes et al.(2015) Escherichia coli   
## 234 362 Berendes et al.(2015) Escherichia coli   
## 235 363 Graham et al. (2003) Cryptosporidium  
## 236 364 Graham et al. (2003) Cryptosporidium  
## 237 365 Graham et al. (2003) Giardia  
## 238 366 Graham et al. (2003) Giardia  
## 239 367 Graham et al. (2003) Fecal coliforms  
## 240 368 Graham et al. (2003) Fecal coliforms  
## 241 369 Robertson et al. (1992) Cryptosporidium  
## 242 370 Robertson et al. (1992) Cryptosporidium  
## 243 371 Robertson et al. (1992) Cryptosporidium  
## bestModel  
## 1 JM2 Model  
## 2 JM2 Model  
## 3 JM2 Model  
## 4 JM2 Model  
## 5 Chick Model  
## 6 Chick Model  
## 7 JM2 Model  
## 8 JM2 Model  
## 9 Chick Model  
## 10 Chick Model  
## 11 Chick Model  
## 12 Chick Model  
## 13 Chick Model  
## 14 JM2 Model  
## 15 JM2 Model  
## 16 JM2 Model  
## 17 JM2 Model  
## 18 JM2 Model  
## 19 JM2 Model  
## 20 Chick Model  
## 21 JM2 Model  
## 22 Chick Model  
## 23 Chick Model  
## 24 Chick Model  
## 25 JM2 Model  
## 26 JM2 Model  
## 27 JM2 Model  
## 28 JM2 Model  
## 29 JM2 Model  
## 30 Chick Model  
## 31 JM2 Model  
## 32 Chick Model  
## 33 Chick Model  
## 34 JM2 Model  
## 35 JM2 Model  
## 36 JM2 Model  
## 37 Chick Model  
## 38 Chick Model  
## 39 Chick Model  
## 40 Chick Model  
## 41 JM2 Model  
## 42 JM2 Model  
## 43 Chick Model  
## 44 JM2 Model  
## 45 JM2 Model  
## 46 JM2 Model  
## 47 JM2 Model  
## 48 JM2 Model  
## 49 Chick Model  
## 50 JM2 Model  
## 51 Chick Model  
## 52 JM2 Model  
## 53 Chick Model  
## 54 Chick Model  
## 55 JM2 Model  
## 56 Chick Model  
## 57 Chick Model  
## 58 JM2 Model  
## 59 JM2 Model  
## 60 Chick Model  
## 61 JM2 Model  
## 62 JM2 Model  
## 63 Chick Model  
## 64 Chick Model  
## 65 JM2 Model  
## 66 Chick Model  
## 67 JM2 Model  
## 68 JM2 Model  
## 69 JM2 Model  
## 70 JM2 Model  
## 71 JM2 Model  
## 72 JM2 Model  
## 73 JM2 Model  
## 74 JM2 Model  
## 75 Chick Model  
## 76 Chick Model  
## 77 Chick Model  
## 78 Chick Model  
## 79 Chick Model  
## 80 Chick Model  
## 81 Chick Model  
## 82 Chick Model  
## 83 Chick Model  
## 84 Chick Model  
## 85 Chick Model  
## 86 Chick Model  
## 87 Chick Model  
## 88 Chick Model  
## 89 Chick Model  
## 90 Chick Model  
## 91 Chick Model  
## 92 Chick Model  
## 93 Chick Model  
## 94 Chick Model  
## 95 Chick Model  
## 96 JM2 Model  
## 97 Chick Model  
## 98 JM2 Model  
## 99 Chick Model  
## 100 Chick Model  
## 101 Chick Model  
## 102 Chick Model  
## 103 Chick Model  
## 104 Chick Model  
## 105 JM2 Model  
## 106 JM2 Model  
## 107 JM2 Model  
## 108 Chick Model  
## 109 Chick Model  
## 110 Chick Model  
## 111 JM2 Model  
## 112 JM2 Model  
## 113 JM2 Model  
## 114 JM2 Model  
## 115 JM2 Model  
## 116 JM2 Model  
## 117 Chick Model  
## 118 Chick Model  
## 119 Chick Model  
## 120 Chick Model  
## 121 Chick Model  
## 122 Chick Model  
## 123 Chick Model  
## 124 Chick Model  
## 125 Chick Model  
## 126 JM2 Model  
## 127 JM2 Model  
## 128 JM2 Model  
## 129 Chick Model  
## 130 Chick Model  
## 131 JM2 Model  
## 132 JM2 Model  
## 133 Chick Model  
## 134 JM2 Model  
## 135 JM2 Model  
## 136 JM2 Model  
## 137 Chick Model  
## 138 Chick Model  
## 139 Chick Model  
## 140 Chick Model  
## 141 Chick Model  
## 142 Chick Model  
## 143 Chick Model  
## 144 Chick Model  
## 145 Chick Model  
## 146 Chick Model  
## 147 Chick Model  
## 148 Chick Model  
## 149 Chick Model  
## 150 Chick Model  
## 151 Chick Model  
## 152 JM2 Model  
## 153 JM2 Model  
## 154 JM2 Model  
## 155 JM2 Model  
## 156 JM2 Model  
## 157 JM2 Model  
## 158 JM2 Model  
## 159 JM2 Model  
## 160 JM2 Model  
## 161 Chick Model  
## 162 JM2 Model  
## 163 Chick Model  
## 164 Chick Model  
## 165 JM2 Model  
## 166 JM2 Model  
## 167 JM2 Model  
## 168 Chick Model  
## 169 JM2 Model  
## 170 JM2 Model  
## 171 JM2 Model  
## 172 JM2 Model  
## 173 Chick Model  
## 174 JM2 Model  
## 175 Chick Model  
## 176 Chick Model  
## 177 JM2 Model  
## 178 Chick Model  
## 179 Chick Model  
## 180 Chick Model  
## 181 Chick Model  
## 182 Chick Model  
## 183 Chick Model  
## 184 Chick Model  
## 185 Chick Model  
## 186 JM2 Model  
## 187 Chick Model  
## 188 Chick Model  
## 189 JM2 Model  
## 190 Chick Model  
## 191 Chick Model  
## 192 JM2 Model  
## 193 Chick Model  
## 194 Chick Model  
## 195 JM2 Model  
## 196 JM2 Model  
## 197 JM2 Model  
## 198 JM2 Model  
## 199 JM2 Model  
## 200 Chick Model  
## 201 JM2 Model  
## 202 JM2 Model  
## 203 JM2 Model  
## 204 Chick Model  
## 205 Chick Model  
## 206 Chick Model  
## 207 JM2 Model  
## 208 Chick Model  
## 209 Chick Model  
## 210 Chick Model  
## 211 Chick Model  
## 212 JM2 Model  
## 213 JM2 Model  
## 214 Chick Model  
## 215 JM2 Model  
## 216 Chick Model  
## 217 Chick Model  
## 218 Chick Model  
## 219 Chick Model  
## 220 JM2 Model  
## 221 JM2 Model  
## 222 Chick Model  
## 223 JM2 Model  
## 224 JM2 Model  
## 225 JM2 Model  
## 226 JM2 Model  
## 227 JM2 Model  
## 228 JM2 Model  
## 229 JM2 Model  
## 230 JM2 Model  
## 231 JM2 Model  
## 232 JM2 Model  
## 233 Chick Model  
## 234 Chick Model  
## 235 Chick Model  
## 236 Chick Model  
## 237 Chick Model  
## 238 Chick Model  
## 239 Chick Model  
## 240 Chick Model  
## 241 JM2 Model  
## 242 JM2 Model  
## 243 JM2 Model

## Produce figures and tables

This last chunk of code uses the data frame kPit (which has all of the results from fitting the models to data from each experiment) to produce figures and tables.

### Table 1

# sample size for k-value calcs based on log linear, then sample size for t99 calcs based on log linear, then the k values, then the estimated T99 values based on the log linear model  
data.frame(rbind(c(paste(length(kPit[kPit$microbial\_group=="Virus",]$k)," (",length(kPit[kPit$microbial\_group=="Virus"&kPit$maxLRV>=1,]$t99),")",sep=""),   
 -quantile(kPit[kPit$microbial\_group=="Virus",]$k,c(0.5,0.95,0.05)),  
 quantile(kPit[kPit$microbial\_group=="Virus"&kPit$maxLRV>=1,]$t99,c(0.5,0.05,0.95))),  
c(paste(length(kPit[kPit$microbial\_group=="Bacteria",]$k)," (",length(kPit[kPit$microbial\_group=="Bacteria"&kPit$maxLRV>=1,]$t99),")",sep=""),   
 -quantile(kPit[kPit$microbial\_group=="Bacteria",]$k,c(0.5,0.95,0.05)),  
 quantile(kPit[kPit$microbial\_group=="Bacteria"&kPit$maxLRV>=1,]$t99,c(0.5,0.05,0.95))),  
c(paste(length(kPit[kPit$microbial\_group=="Protozoa",]$k)," (",length(kPit[kPit$microbial\_group=="Protozoa"&kPit$maxLRV>=1,]$t99),")",sep=""),   
 -quantile(kPit[kPit$microbial\_group=="Protozoa",]$k,c(0.5,0.95,0.05)),  
 quantile(kPit[kPit$microbial\_group=="Protozoa"&kPit$maxLRV>=1,]$t99,c(0.5,0.05,0.95))),  
c(paste(length(kPit[kPit$microbial\_group=="Helminth",]$k)," (",length(kPit[kPit$microbial\_group=="Helminth"&kPit$maxLRV>=1,]$t99),")",sep=""),   
 -quantile(kPit[kPit$microbial\_group=="Helminth",]$k,c(0.5,0.95,0.05)),  
 quantile(kPit[kPit$microbial\_group=="Helminth"&kPit$maxLRV>=1,]$t99,c(0.5,0.05,0.95)))  
))

## V1 X50. X95. X5. X50..1 X5..1  
## 1 32 (31) 0.18885 0.02547 16.800765 29.8092087720436 0.940083546180488  
## 2 147 (120) 0.9313 0.01164 242.79872 0.90902445630565 0.00242002305642991  
## 3 8 (4) 0.00945 0.001605 0.03501 269.526297347728 125.84936663375  
## 4 56 (23) 0.04165 5e-05 0.52 44.3065754116127 6.62139246819842  
## X95..1  
## 1 134.199331043184  
## 2 72.4674672473823  
## 3 320.697383842882  
## 4 290.930499952965

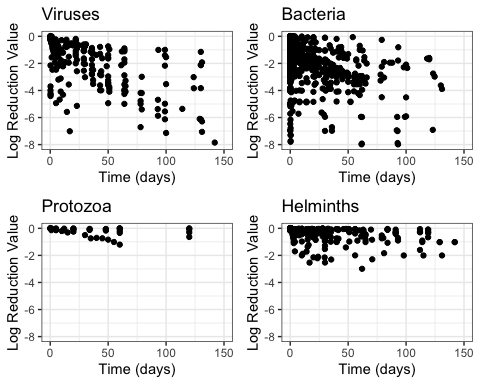
### Table 2

data.frame(rbind(  
 c(length(kPit[kPit$microbial\_group=="Virus"&kPit$num>3&kPit$maxLRV>=1,]$t99jm),  
 quantile(kPit[kPit$microbial\_group=="Virus"&kPit$num>3&kPit$maxLRV>=1,]$k1,c(0.5,0.05,0.95),na.rm=T),  
 quantile(kPit[kPit$microbial\_group=="Virus"&kPit$num>3&kPit$maxLRV>=1,]$k2,c(0.5,0.05,0.95),na.rm=T),  
 quantile(kPit[kPit$microbial\_group=="Virus"&kPit$num>3&kPit$maxLRV>=1,]$t99jm,c(0.5,0.05,0.95),na.rm=T)),  
 c(length(kPit[kPit$microbial\_group=="Bacteria"&kPit$num>3&kPit$maxLRV>=1,]$t99jm),  
 quantile(kPit[kPit$microbial\_group=="Bacteria"&kPit$num>3&kPit$maxLRV>=1,]$k1,c(0.5,0.05,0.95),na.rm=T),  
 quantile(kPit[kPit$microbial\_group=="Bacteria"&kPit$num>3&kPit$maxLRV>=1,]$k2,c(0.5,0.05,0.95),na.rm=T),  
 quantile(kPit[kPit$microbial\_group=="Bacteria"&kPit$num>3&kPit$maxLRV>=1,]$t99jm,c(0.5,0.05,0.95),na.rm=T)),  
 c(length(kPit[kPit$microbial\_group=="Protozoa"&kPit$num>3&kPit$maxLRV>=1,]$t99jm),  
 quantile(kPit[kPit$microbial\_group=="Protozoa"&kPit$num>3&kPit$maxLRV>=1,]$k1,c(0.5,0.05,0.95),na.rm=T),  
 quantile(kPit[kPit$microbial\_group=="Protozoa"&kPit$num>3&kPit$maxLRV>=1,]$k2,c(0.5,0.05,0.95),na.rm=T),  
 quantile(kPit[kPit$microbial\_group=="Protozoa"&kPit$num>3&kPit$maxLRV>=1,]$t99jm,c(0.5,0.05,0.95),na.rm=T)),  
 c(length(kPit[kPit$microbial\_group=="Helminth"&kPit$num>3&kPit$maxLRV>=1,]$t99jm),  
 quantile(kPit[kPit$microbial\_group=="Helminth"&kPit$num>3&kPit$maxLRV>=1,]$k1,c(0.5,0.05,0.95),na.rm=T),  
 quantile(kPit[kPit$microbial\_group=="Helminth"&kPit$num>3&kPit$maxLRV>=1,]$k2,c(0.5,0.05,0.95),na.rm=T),  
 quantile(kPit[kPit$microbial\_group=="Helminth"&kPit$num>3&kPit$maxLRV>=1,]$t99jm,c(0.5,0.05,0.95),na.rm=T))  
))

## V1 X50. X5. X95. X50..1 X5..1 X95..1 X50..2 X5..2  
## 1 31 -7.190 -31.0500 17.725 4.920 1.485 17.870 29.367170 0.9675115  
## 2 76 0.875 -78.3400 215.855 3.625 0.450 215.855 8.677702 0.1009299  
## 3 3 -13.360 -13.6210 -6.547 3.130 2.104 3.310 220.666726 190.4310660  
## 4 18 -8.675 -58.1235 -1.847 3.830 1.389 17.626 61.532533 12.3873031  
## X95..2  
## 1 187.9810  
## 2 158.3257  
## 3 328.6808  
## 4 457.6740

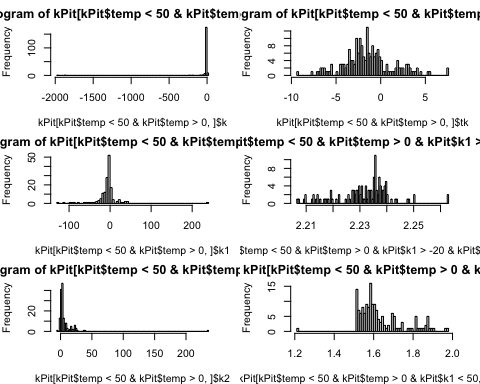
### Figure 2

persistExc<-persist[!persist$experiment\_id %in% kPit$experiment,]  
persistInc<-persist[persist$experiment\_id %in% kPit$experiment,]  
  
plotFun<-function(persistanceData){  
 fitV <- lm(-log10\_reduction ~ time\_days, data = persistanceData[persistanceData$microbial\_group=="Virus" & persistanceData$matrix=="Fecal sludge" & persistanceData$time\_days<200,])  
 new\_df <- persistanceData[persistanceData$microbial\_group=="Virus" & persistanceData$matrix=="Fecal sludge" & persistanceData$time\_days<200,]   
 vi <- ggplot(new\_df, aes(time\_days, -log10\_reduction)) +  
 geom\_point() +  
 ylim(-8,0) + ylab("Log Reduction Value") +   
 xlim(0,150) + xlab("Time (days)") +  
 ggtitle("Viruses") +  
 theme(plot.title = element\_text(hjust=0.5)) +  
 theme\_bw() #+ geom\_smooth(method=lm, se=TRUE)  
 fitB <- lm(-log10\_reduction ~ time\_days, data = persistanceData[persistanceData$microbial\_group=="Bacteria" & persistanceData$matrix=="Fecal sludge" & persistanceData$time\_days<200,])  
 new\_df <- persistanceData[persistanceData$microbial\_group=="Bacteria" & persistanceData$matrix=="Fecal sludge" & persistanceData$time\_days<200,]  
 ba <- ggplot(new\_df, aes(time\_days, -log10\_reduction))+  
 geom\_point() +  
 ylim(-8,0) + ylab("Log Reduction Value") +   
 xlim(0,150) + xlab("Time (days)") +  
 ggtitle("Bacteria") +  
 theme(plot.title = element\_text(hjust=0.5))+  
 theme\_bw()  
 fitH <- lm(-log10\_reduction ~ time\_days, data = persistanceData[persistanceData$microbial\_group=="Helminth" & persistanceData$matrix=="Fecal sludge" & persistanceData$time\_days<200,])  
 new\_df <- persistanceData[persistanceData$microbial\_group=="Helminth" & persistanceData$matrix=="Fecal sludge" & persistanceData$time\_days<200,]  
 he <- ggplot(new\_df, aes(time\_days, -log10\_reduction))+  
 geom\_point() +  
 ylim(-8,0) + ylab("Log Reduction Value") +   
 xlim(0,150) + xlab("Time (days)") +  
 ggtitle("Helminths") +  
 theme(plot.title = element\_text(hjust=0.5))+  
 theme\_bw()  
 fitP <- lm(-log10\_reduction ~ time\_days, data = persistanceData[persistanceData$microbial\_group=="Protozoa" & persistanceData$matrix=="Fecal sludge" & persistanceData$time\_days<200,])  
 new\_df <- persistanceData[persistanceData$microbial\_group=="Protozoa" & persistanceData$matrix=="Fecal sludge" & persistanceData$time\_days<200,]  
 pr <- ggplot(new\_df, aes(time\_days, -log10\_reduction))+  
 geom\_point() +  
 ylim(-8,0) + ylab("Log Reduction Value") +   
 xlim(0,150) + xlab("Time (days)") +  
 ggtitle("Protozoa") +  
 theme(plot.title = element\_text(hjust=0.5))+  
 theme\_bw()  
grid.arrange(vi,ba,pr,he,nrow = 2,widths = c(2,2))  
}  
plotFun(persistInc) # Plots Figure 2



### Figure S4

kPit$tk<-log(-kPit$k)  
kPit$tk1<-sqrt(log(kPit$k1+150))  
kPit$tk2<-sqrt(log(kPit$k2+10))  
  
par(mfrow=c(3,2),mar=c(4,4,2,2))  
hist(kPit[kPit$temp<50&kPit$temp>0,]$k,breaks=100)  
hist(kPit[kPit$temp<50&kPit$temp>0,]$tk,breaks=100)  
hist(kPit[kPit$temp<50&kPit$temp>0,]$k1,breaks=100)  
hist(kPit[kPit$temp<50&kPit$temp>0&kPit$k1>-20&kPit$k1<20,]$tk1,breaks=100)  
hist(kPit[kPit$temp<50&kPit$temp>0,]$k2,breaks=100)  
hist(kPit[kPit$temp<50&kPit$temp>0&kPit$k1<50,]$tk2,breaks=100)



### Figure S6

myData<-kPit[kPit$temp<50&kPit$temp>0,c("tk","k1","tk2","temp","pH","moisture")]  
names(myData)<-c("ln(k) (1/days)","k1","tk2","Temperature (°C)","pH","moisture")  
par(mfrow=c(1,2))  
chart.Correlation(kPit[,c("tk","k1","tk2","temp","pH","moisture")],histogram=TRUE)

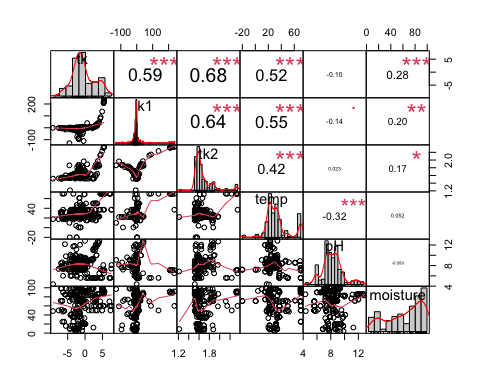


chart.Correlation(myData,histogram=TRUE)

