Rangewide Analysis

Andres N. Rosales

2025-01-24

Table of Contents

# Candidate Models

**Prediction 1**

Breeding regions declining in trend will be correlated with low landscape heterogeneity and simplified agricultural landscapes.

**Prediction 2**

Greater landscape heterogeneity will lower the negative effects of landscape simplification on breeding trends. “Is the effect of landscape diversity on Upland Sandpiper population dynamics dependent on the degree of simplification?”

**Prediction 3**

Forage agriculture and grasslands are associated with positive trends and high abundance. “How are trends in forage agriculture and grasslands associated with trends and abundance of Upland Sandpiper?”

“Does configuration of forage and grassland habitats influence their association with trends and abundance

**Prediction 4**

Trends decline or become more variable with increasing distance from the core breeding range. Conversely, trends will be more stable near the core breeding region.

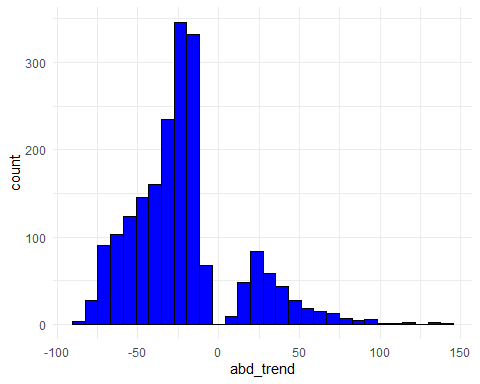
**Prediction 5**

Geographical variation in abundance is driven by marginal habitat at the periphery.

# Response Distribution

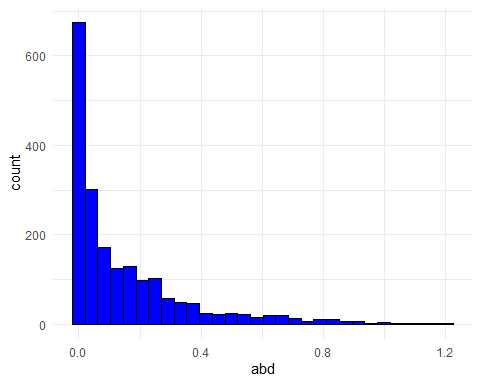
Distribution of trend response variable.

ggplot(upsabuff, aes(x = abd\_trend)) +  
 geom\_histogram(bins = 30, fill = "blue", color = "black") +  
 theme\_minimal()

 Almost follows a normal distribution, but without trends that are zero (stable) we observe biomodal. I could possibly look into adding a categorical variable indicating increase or decrease. I need to explore this more…

#add categorical variable to each location indicating increase or decrease  
upsa\_trd$trend\_dir <- ifelse(upsa\_trd$abd\_trend >0, "Increase", "Decrease")

#Distribution of abundance response variable.   
ggplot(upsabuff, aes(x = abd)) +  
 geom\_histogram(bins = 30, fill = "blue", color = "black") +  
 theme\_minimal()

 Distribution of abundance estimates.

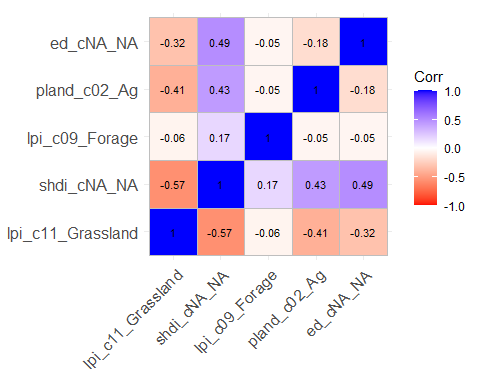
# Predictor relationships

Correlation matric before aggregating predictors.

library(ggcorrplot)

## Warning: package 'ggcorrplot' was built under R version 4.4.2

# List of predictors to include  
predictors <- c(  
 "lpi\_c11\_Grassland", "shdi\_cNA\_NA", "lpi\_c09\_Forage", "pland\_c02\_Ag", "ed\_cNA\_NA"  
)  
  
# Filter dataset  
filtered\_data <- all\_lsm %>%  
 dplyr::select(all\_of(predictors))  
  
# Compute correlation matrix  
cor\_matrix <- cor(filtered\_data, use = "complete.obs", method = "pearson")  
  
# Visualize the correlation matrix  
ggcorrplot(cor\_matrix, lab = TRUE, lab\_size = 3, colors = c("red", "white", "blue"))



I still need to determine a pearsons correlation coefficient threshold to report. Some literature has used 0.7 and I will go with that for now. I still need to develop a better backing for this and if this is something I will continue to use going forward…

## Shannon Diversity Index

I want to use the slope of a linear regression for SHDI ~ year for every location as a proxy for change in SHDI. I will need to confirm SHDI changes in linearly. fit lm and gam and compare fit for slope of SHDI as explanatory predictor

shdi\_lmm <- lmer(shdi\_cNA\_NA ~ year + (1|srd\_id), data = all\_lsm)  
summary(shdi\_lmm)

## Linear mixed model fit by REML ['lmerMod']  
## Formula: shdi\_cNA\_NA ~ year + (1 | srd\_id)  
## Data: all\_lsm  
##   
## REML criterion at convergence: -34837.1  
##   
## Scaled residuals:   
## Min 1Q Median 3Q Max   
## -14.0486 -0.4342 -0.0162 0.4080 10.1691   
##   
## Random effects:  
## Groups Name Variance Std.Dev.  
## srd\_id (Intercept) 0.149392 0.38651   
## Residual 0.007152 0.08457   
## Number of obs: 21681, groups: srd\_id, 1971  
##   
## Fixed effects:  
## Estimate Std. Error t value  
## (Intercept) -8.9275736 0.3664480 -24.36  
## year 0.0051451 0.0001816 28.33  
##   
## Correlation of Fixed Effects:  
## (Intr)  
## year -1.000

shdi\_gamm <- gam(shdi\_cNA\_NA ~ s(year, k=10) + srd\_id, method="REML", data = all\_lsm)  
summary(shdi\_gamm)

##   
## Family: gaussian   
## Link function: identity   
##   
## Formula:  
## shdi\_cNA\_NA ~ s(year, k = 10) + srd\_id  
##   
## Parametric coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 2.443e+00 2.672e-02 91.42 <2e-16 \*\*\*  
## srd\_id -3.503e-06 9.382e-08 -37.33 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Approximate significance of smooth terms:  
## edf Ref.df F p-value   
## s(year) 4.291 5.282 12.09 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## R-sq.(adj) = 0.063 Deviance explained = 6.32%  
## -REML = 9997.3 Scale est. = 0.14687 n = 21681

AIC(shdi\_lmm, shdi\_gamm)

## df AIC  
## shdi\_lmm 4.000000 -34829.10  
## shdi\_gamm 8.281729 19949.43

Big difference in AIC values when comparing GAM and lm model. the relationships of each location likely follow a linear relationship over years, allowing it to be representative of the change in diversity over time.

Calculate change in SHDI using linear regression. THis will be used as a predictor representing the change in diversity over the trend period

#Slope and intercept  
shdi\_slopes <- all\_lsm %>%  
 group\_by(srd\_id) %>%  
 summarize(  
 shdi\_slope = coef(lm(shdi\_cNA\_NA ~ year))[2],   
 shdi\_intercept = coef(lm(shdi\_cNA\_NA ~ year))[1]   
 )  
  
upsa\_trd <- shdi\_slopes %>%  
 left\_join(upsa\_trd, by = c("srd\_id"))

add mean and standard deviation of shdi for each location

#Mean  
mean\_shdi <- all\_lsm %>%  
 group\_by(srd\_id) %>%  
 summarize(mean\_shdi = mean(shdi\_cNA\_NA, na.rm = TRUE))  
upsa\_trd <- mean\_shdi %>%  
 left\_join(upsa\_trd, by = "srd\_id")  
#Standard Dev  
sd\_shdi <- all\_lsm %>%  
 group\_by(srd\_id) %>%  
 summarize(sd\_shdi = sd(shdi\_cNA\_NA, na.rm = TRUE))  
  
upsa\_trd <- sd\_shdi %>%  
 left\_join(upsa\_trd, by = "srd\_id")

ggplot(upsa\_trd, aes(x = shdi\_slope, y = abd\_trend)) +  
 geom\_point(color = "blue", alpha = 0.3, size = 2) +   
 geom\_smooth(method = "lm", color = "red", se = TRUE, linetype = "dashed") +   
 labs(  
 title = "Relationship Between SHDI Slope and Trend",  
 x = "Slope of SHDI (Change over Time)",  
 y = "Trend"  
 ) +  
 theme\_minimal() +  
 theme(  
 plot.title = element\_text(hjust = 0.5, size = 16),  
 axis.title = element\_text(size = 14),  
 axis.text = element\_text(size = 12)  
 )

## `geom\_smooth()` using formula = 'y ~ x'

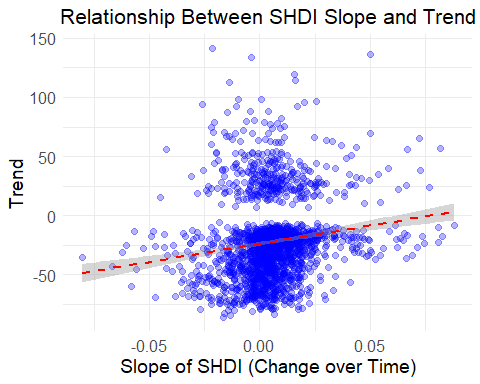
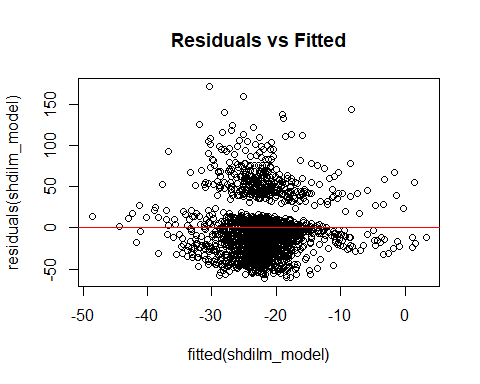


Figure reveals mostly linear with some deviation along the edges. Test lm and gam model to determine best fit.

shdilm\_model <- lm(abd\_trend ~ shdi\_slope, data = upsa\_trd)  
plot(fitted(shdilm\_model), residuals(shdilm\_model), main = "Residuals vs Fitted")  
abline(h = 0, col = "red")



shdigam\_model <- gam(abd\_trend ~ s(shdi\_slope), data = upsa\_trd)  
summary(shdilm\_model)

##   
## Call:  
## lm(formula = abd\_trend ~ shdi\_slope, data = upsa\_trd)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -61.388 -20.679 -3.738 7.770 172.118   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -23.7973 0.7742 -30.738 < 2e-16 \*\*\*  
## shdi\_slope 308.4663 44.9874 6.857 9.4e-12 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 32.8 on 1969 degrees of freedom  
## Multiple R-squared: 0.02332, Adjusted R-squared: 0.02282   
## F-statistic: 47.01 on 1 and 1969 DF, p-value: 9.396e-12

summary(shdigam\_model)

##   
## Family: gaussian   
## Link function: identity   
##   
## Formula:  
## abd\_trend ~ s(shdi\_slope)  
##   
## Parametric coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -22.2102 0.7379 -30.1 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Approximate significance of smooth terms:  
## edf Ref.df F p-value   
## s(shdi\_slope) 2.633 3.388 15.13 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## R-sq.(adj) = 0.0252 Deviance explained = 2.65%  
## GCV = 1075.1 Scale est. = 1073.1 n = 1971

# Compare lm and gam  
AIC(shdilm\_model, shdigam\_model)

## df AIC  
## shdilm\_model 3.00000 19356.62  
## shdigam\_model 4.63294 19353.40

Slight improvement of fit with the GAM model. However, since it is small it might be ok to assume linear relationship.

## Landscape Simplification

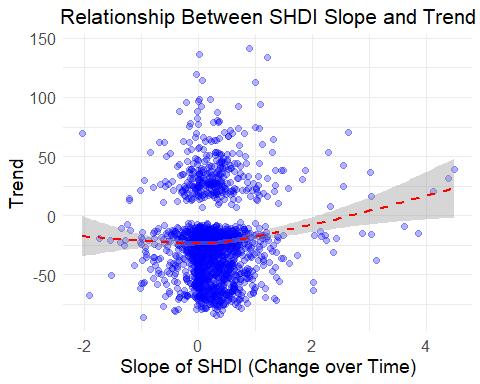
calculate slope of landscape simplification over 11 years

simp\_slopes <- all\_lsm %>%  
 group\_by(srd\_id) %>%  
 summarize(  
 simp\_slope = coef(lm(pland\_c02\_Ag ~ year))[2],   
 simp\_intercept = coef(lm(pland\_c02\_Ag ~ year))[1]   
 )  
  
upsa\_trd <- simp\_slopes %>%  
 left\_join(upsa\_trd, by = c("srd\_id"))

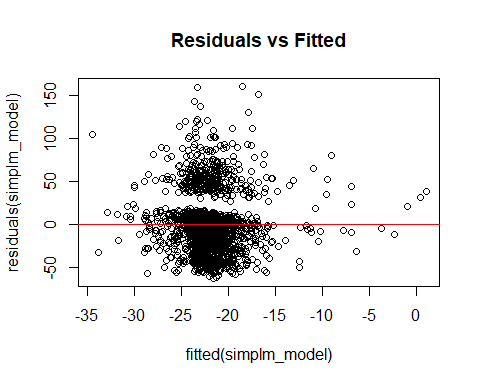
plot simplification slope against trend

ggplot(upsa\_trd, aes(x = simp\_slope, y = abd\_trend)) +  
 geom\_point(color = "blue", alpha = 0.3, size = 2) +   
 geom\_smooth(method = "gam", color = "red", se = TRUE, linetype = "dashed") +   
 labs(  
 title = "Relationship Between SHDI Slope and Trend",  
 x = "Slope of SHDI (Change over Time)",  
 y = "Trend"  
 ) +  
 theme\_minimal() +  
 theme(  
 plot.title = element\_text(hjust = 0.5, size = 16),  
 axis.title = element\_text(size = 14),  
 axis.text = element\_text(size = 12)  
 )

## `geom\_smooth()` using formula = 'y ~ s(x, bs = "cs")'



simplm\_model <- lm(abd\_trend ~ simp\_slope, data = upsa\_trd)  
plot(fitted(simplm\_model), residuals(simplm\_model), main = "Residuals vs Fitted")  
abline(h = 0, col = "red")



simpgam\_model <- gam(abd\_trend ~ s(simp\_slope), data = upsa\_trd)  
summary(simplm\_model)

##   
## Call:  
## lm(formula = abd\_trend ~ simp\_slope, data = upsa\_trd)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -62.918 -21.016 -2.607 8.113 160.319   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -23.4372 0.8155 -28.740 < 2e-16 \*\*\*  
## simp\_slope 5.4440 1.4719 3.699 0.000223 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 33.07 on 1969 degrees of freedom  
## Multiple R-squared: 0.0069, Adjusted R-squared: 0.006395   
## F-statistic: 13.68 on 1 and 1969 DF, p-value: 0.0002227

summary(simpgam\_model)

##   
## Family: gaussian   
## Link function: identity   
##   
## Formula:  
## abd\_trend ~ s(simp\_slope)  
##   
## Parametric coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -22.2102 0.7428 -29.9 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Approximate significance of smooth terms:  
## edf Ref.df F p-value   
## s(simp\_slope) 2.977 3.812 6.359 7.28e-05 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## R-sq.(adj) = 0.0122 Deviance explained = 1.37%  
## GCV = 1089.7 Scale est. = 1087.5 n = 1971

# Compare lm and gam  
AIC(simplm\_model, simpgam\_model)

## df AIC  
## simplm\_model 3.00000 19389.49  
## simpgam\_model 4.97683 19379.94

## LPI Grassland

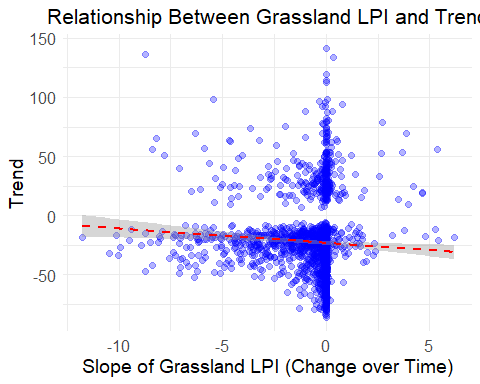
calculate the slope and average LPI for each location

#Slope  
GLPI\_slopes <- all\_lsm %>%  
 group\_by(srd\_id) %>%  
 summarize(  
 GLPI\_slope = coef(lm(lpi\_c11\_Grassland ~ year))[2],   
 GLPI\_intercept = coef(lm(lpi\_c11\_Grassland ~ year))[1]   
 )  
upsa\_trd <- GLPI\_slopes %>%  
 left\_join(upsa\_trd, by = c("srd\_id"))  
  
#Mean  
mean\_GLPI <- all\_lsm %>%  
 group\_by(srd\_id) %>%  
 summarize(mean\_GLPI = mean(lpi\_c11\_Grassland, na.rm = TRUE))  
upsa\_trd <- mean\_GLPI %>%  
 left\_join(upsa\_trd, by = "srd\_id")

Plot trend ~ slope and trend ~ Mean

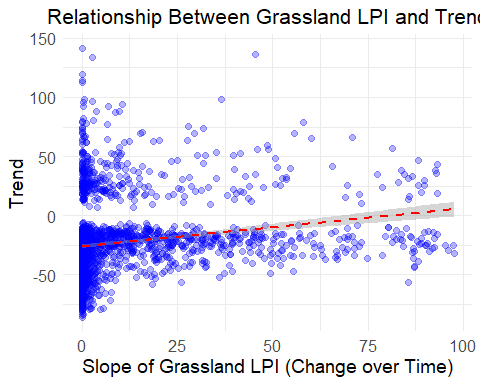
ggplot(upsa\_trd, aes(x = GLPI\_slope, y = abd\_trend)) +  
 geom\_point(color = "blue", alpha = 0.3, size = 2) +   
 geom\_smooth(method = "lm", color = "red", se = TRUE, linetype = "dashed") +   
 labs(  
 title = "Relationship Between Grassland LPI and Trend",  
 x = "Slope of Grassland LPI (Change over Time)",  
 y = "Trend"  
 ) +  
 theme\_minimal() +  
 theme(  
 plot.title = element\_text(hjust = 0.5, size = 16),  
 axis.title = element\_text(size = 14),  
 axis.text = element\_text(size = 12)  
 )

## `geom\_smooth()` using formula = 'y ~ x'



ggplot(upsa\_trd, aes(x = mean\_GLPI, y = abd\_trend)) +  
 geom\_point(color = "blue", alpha = 0.3, size = 2) +   
 geom\_smooth(method = "lm", color = "red", se = TRUE, linetype = "dashed") +   
 labs(  
 title = "Relationship Between Grassland LPI and Trend",  
 x = "Slope of Grassland LPI (Change over Time)",  
 y = "Trend"  
 ) +  
 theme\_minimal() +  
 theme(  
 plot.title = element\_text(hjust = 0.5, size = 16),  
 axis.title = element\_text(size = 14),  
 axis.text = element\_text(size = 12)  
 )

## `geom\_smooth()` using formula = 'y ~ x'



## LPI Forage

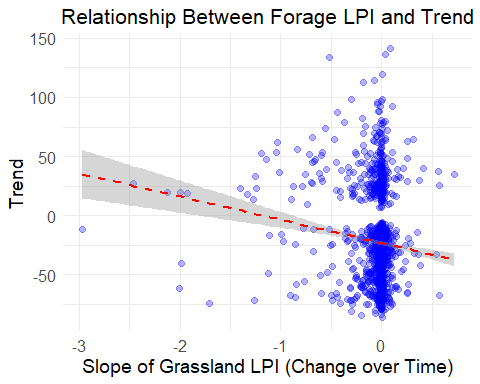
calculate the slope and average LPI for each location

#Slope  
FLPI\_slopes <- all\_lsm %>%  
 group\_by(srd\_id) %>%  
 summarize(  
 FLPI\_slope = coef(lm(lpi\_c09\_Forage ~ year))[2],   
 FLPI\_intercept = coef(lm(lpi\_c09\_Forage ~ year))[1]   
 )  
upsa\_trd <- FLPI\_slopes %>%  
 left\_join(upsa\_trd, by = c("srd\_id"))  
  
#Mean  
mean\_FLPI <- all\_lsm %>%  
 group\_by(srd\_id) %>%  
 summarize(mean\_FLPI = mean(lpi\_c09\_Forage, na.rm = TRUE))  
upsa\_trd <- mean\_FLPI %>%  
 left\_join(upsa\_trd, by = "srd\_id")

Plot trend ~ slope and trend ~ Mean

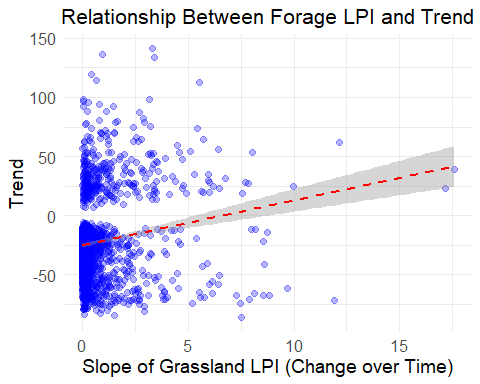
ggplot(upsa\_trd, aes(x = FLPI\_slope, y = abd\_trend)) +  
 geom\_point(color = "blue", alpha = 0.3, size = 2) +   
 geom\_smooth(method = "lm", color = "red", se = TRUE, linetype = "dashed") +   
 labs(  
 title = "Relationship Between Forage LPI and Trend",  
 x = "Slope of Grassland LPI (Change over Time)",  
 y = "Trend"  
 ) +  
 theme\_minimal() +  
 theme(  
 plot.title = element\_text(hjust = 0.5, size = 16),  
 axis.title = element\_text(size = 14),  
 axis.text = element\_text(size = 12)  
 )

## `geom\_smooth()` using formula = 'y ~ x'



ggplot(upsa\_trd, aes(x = mean\_FLPI, y = abd\_trend)) +  
 geom\_point(color = "blue", alpha = 0.3, size = 2) +   
 geom\_smooth(method = "lm", color = "red", se = TRUE, linetype = "dashed") +   
 labs(  
 title = "Relationship Between Forage LPI and Trend",  
 x = "Slope of Grassland LPI (Change over Time)",  
 y = "Trend"  
 ) +  
 theme\_minimal() +  
 theme(  
 plot.title = element\_text(hjust = 0.5, size = 16),  
 axis.title = element\_text(size = 14),  
 axis.text = element\_text(size = 12)  
 )

## `geom\_smooth()` using formula = 'y ~ x'



## Edge Density

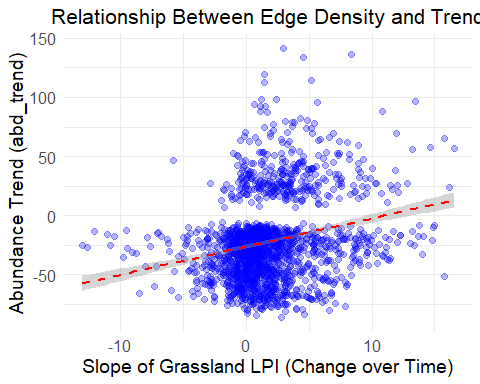
calcaulte the slope and average LPI for each location

#Slope  
ED\_slopes <- all\_lsm %>%  
 group\_by(srd\_id) %>%  
 summarize(  
 ED\_slope = coef(lm(ed\_cNA\_NA ~ year))[2],   
 ED\_intercept = coef(lm(ed\_cNA\_NA ~ year))[1]   
 )  
upsa\_trd <- ED\_slopes %>%  
 left\_join(upsa\_trd, by = c("srd\_id"))  
  
#Mean  
mean\_ED <- all\_lsm %>%  
 group\_by(srd\_id) %>%  
 summarize(mean\_ED = mean(ed\_cNA\_NA, na.rm = TRUE))  
upsa\_trd <- mean\_ED %>%  
 left\_join(upsa\_trd, by = "srd\_id")

Plot trend ~ slope and trend ~ Mean

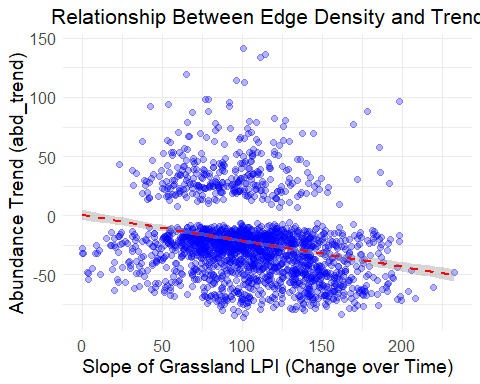
ggplot(upsa\_trd, aes(x = ED\_slope, y = abd\_trend)) +  
 geom\_point(color = "blue", alpha = 0.3, size = 2) +   
 geom\_smooth(method = "lm", color = "red", se = TRUE, linetype = "dashed") +   
 labs(  
 title = "Relationship Between Edge Density and Trend",  
 x = "Slope of Grassland LPI (Change over Time)",  
 y = "Abundance Trend (abd\_trend)"  
 ) +  
 theme\_minimal() +  
 theme(  
 plot.title = element\_text(hjust = 0.5, size = 16),  
 axis.title = element\_text(size = 14),  
 axis.text = element\_text(size = 12)  
 )

## `geom\_smooth()` using formula = 'y ~ x'

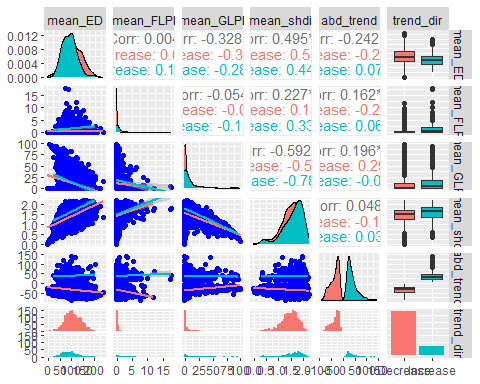


ggplot(upsa\_trd, aes(x = mean\_ED, y = abd\_trend)) +  
 geom\_point(color = "blue", alpha = 0.3, size = 2) +   
 geom\_smooth(method = "lm", color = "red", se = TRUE, linetype = "dashed") +   
 labs(  
 title = "Relationship Between Edge Density and Trend",  
 x = "Slope of Grassland LPI (Change over Time)",  
 y = "Abundance Trend (abd\_trend)"  
 ) +  
 theme\_minimal() +  
 theme(  
 plot.title = element\_text(hjust = 0.5, size = 16),  
 axis.title = element\_text(size = 14),  
 axis.text = element\_text(size = 12)  
 )

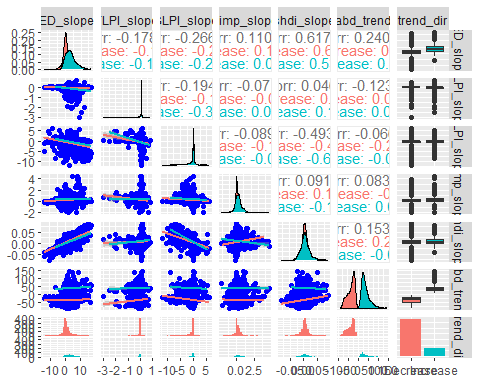
## `geom\_smooth()` using formula = 'y ~ x'

 ### Correlation matrix matrix of relationships / distribution of response and explanatory variables after aggregating

upsa\_trd$trend\_dir <- ifelse(upsa\_trd$abd\_trend >0, "Increase", "Decrease")  
  
  
agg\_mean\_predictors <- upsa\_trd[,c("mean\_ED","mean\_FLPI","mean\_GLPI", "mean\_shdi", "abd\_trend", "trend\_dir")]  
  
agg\_slope\_predictors <- upsa\_trd[,c("ED\_slope","FLPI\_slope", "GLPI\_slope", "simp\_slope", "shdi\_slope", "abd\_trend", "trend\_dir")]  
  
#aggregated\_predictors <- upsa\_trd[,c("mean\_ED","ED\_slope","mean\_FLPI","FLPI\_slope","mean\_GLPI", "GLPI\_slope", "simp\_slope", "mean\_shdi", "shdi\_slope", "abd\_trend", "trend\_dir")]  
  
#Average   
ggpairs(agg\_mean\_predictors, ggplot2::aes(colour = trend\_dir), lower = list(continuous = wrap("smooth", colour = "blue")))



#Slope  
ggpairs(agg\_slope\_predictors, ggplot2::aes(colour = trend\_dir), lower = list(continuous = wrap("smooth", colour = "blue")))



names(upsa\_trd)

## [1] "srd\_id" "mean\_ED" "ED\_slope" "ED\_intercept"   
## [5] "mean\_FLPI" "FLPI\_slope" "FLPI\_intercept" "mean\_GLPI"   
## [9] "GLPI\_slope" "GLPI\_intercept" "simp\_slope" "simp\_intercept"   
## [13] "sd\_shdi" "mean\_shdi" "shdi\_slope" "shdi\_intercept"   
## [17] "longitude" "latitude" "abd" "abd\_ppy"   
## [21] "abd\_ppy\_lower" "abd\_ppy\_upper" "abd\_ppy\_nonzero" "abd\_trend"   
## [25] "abd\_trend\_lower" "abd\_trend\_upper" "trend\_dir"

# Model Runs

Current work in progress…

I have the data frame and predictors set up to be able to run models.

I need to determine which predictors I will use in candidate models (mean vs slope). Also, I need to determine a strategy for how I will model.(How can I include abundance, spatial dependencies?, error structure?)