Compare Amphibian Dermal Exposure Data with Proposed Models

Purucker ST, Snyder MN, Glinski DA, Van Meter RJ, Garber K, Cyterski MJ, Sinnathamby S, Henderson WM

Overview

A comparison between the Terrestrial Investigation Model (TIM) and the collected amphibian dermal exposure data (introduced in 00_amphib_data_combine.Rmd) will be conducted. TIM is a probabilistic model simulating multimedia pesticide exposure from food, dermal contact, drinking water, or inhalation exposure routes.

Collected amphibian dermal exposure data is from 9 studies:

- Henson-Ramsey et al. 2008 data consists of exposure to malathion for Tiger salamanders.
- Van Meter et al. 2014/2015 data includes exposure with 5 active ingredients (imidacloprid, pendimethalin, atrazine, fipronil, and triadimefon) and 9 amphibian species in the terrestrial metamorph stage (American toad, Barking treefrog, Cricket frog, Fowler's toad, Gray treefrog, Green treefrog, Leopard frog, Mole salamander, Narrowmouth toad).
- Van Meter et al. 2016 data includes exposure to the same 5 pesticides on American toads. Van Meter et al. 2018 data consisted of exposure to a single pesticide or pesticide mixtures (atrazine, metolachlor, 2,4-D, malathion, propiconazole) on juvenile green frogs. *Van Meter et al. 2018 data that looked at soil applications, including mixtures.
- Van Meter et al 2019 data that jointly examined exposures of pesticides and fertilizers.
- Unpublished Van Meter 2021 data that uses the same techniques as Van Meter et al. 2019
- Glinski et al. 2018a (Dehydration) data includes exposure to 5 pesticides (atrazine, triadimefon, metolachlor, chlorothalonil, imidacloprid) on 2 amphibian species (Southern leopard frogs, Fowler's toads).
- Glinski et al. 2018b (Metabolites) data consists of exposure to 3 pesticides (atrazine, triadimefon, fipronil) on Fowler's toads.
- Glinski et al. 2019 (Biomarkers) data includes exposure to single, double, or triple pesticide mixtures (bifenthrin, metolachlor, triadimefon) on Southern leopard frogs.

- Glinski et al. 2021 (Dermal Routes) data includes exposure to bifenthrin, chlorpyrifos, and trifloxystrobin on Leopard frogs.

Collated Data Set

Data Set Dimensions, Column Names, and Summary:

```
## [1] 1158   13
```

```
##  [1] "X"               "app_rate_g_cm2"  "application"     "body_weight_g"
##  [5] "chemical"        "exp_duration"    "formulation"     "sample_id"
##  [9] "soil_conc_ugg"   "soil_type"       "source"          "species"
## [13] "tissue_conc_ugg"
```

```
##  [1] "imidacloprid"    "pendimethalin"   "atrazine"        "fipronil"
##  [5] "triadimefon"     "chlorothalonil"  "metolachlor"     "malathion"
##  [9] "bifenthrin"      "dt"              "propiconazole"   "chlorpyrifos"
## [13] "trifloxystrobin" "alachlor"
```

```
##  [1] "Barking treefrog"      "Cricket frog"          "Fowlers toad"
##  [4] "Gray treefrog"         "Green treefrog"        "Leopard frog"
##  [7] "Mole salamander"       "Narrowmouth toad"      "American toad"
## [10] "Ambystoma_tigrinum"    "Rana_clamitans"        "Southern Leopard Frog"
```

```
##        X          app_rate_g_cm2      application        body_weight_g
##  Min.   :   1.0   Min.   :2.870e-07   Length:1158        Min.   : 0.1879
##  1st Qu.: 290.2   1st Qu.:2.870e-06   Class :character   1st Qu.: 1.4628
##  Median : 579.5   Median :2.280e-05   Mode  :character   Median : 2.2345
##  Mean   : 579.5   Mean   :1.778e-05                      Mean   : 3.3645
##  3rd Qu.: 868.8   3rd Qu.:3.062e-05                      3rd Qu.: 3.1265
##  Max.   :1158.0   Max.   :6.980e-05                      Max.   :50.9200
##
##    chemical          exp_duration     formulation       sample_id
##  Length:1158        Min.   : 2.000   Min.   :0.00000   Length:1158
##  Class :character   1st Qu.: 8.000   1st Qu.:0.00000   Class :character
##  Mode  :character   Median : 8.000   Median :0.00000   Mode  :character
##                     Mean   : 8.767   Mean   :0.03133
##                     3rd Qu.: 8.000   3rd Qu.:0.00000
##                     Max.   :48.000   Max.   :1.00000
##                                      NA's   :9
##  soil_conc_ugg        soil_type            source            species
##  Min.   :   0.1125   Length:1158        Length:1158        Length:1158
##  1st Qu.:   2.2855   Class :character   Class :character   Class :character
##  Median :   7.0111   Mode  :character   Mode  :character   Mode  :character
##  Mean   :  14.9202
##  3rd Qu.:  16.6984
##  Max.   :1283.5953
##  NA's   :9
##  tissue_conc_ugg
##  Min.   : 0.00054
##  1st Qu.: 0.20029
##  Median : 0.61249
##  Mean   : 2.44180
##  3rd Qu.: 2.07634
##  Max.   :72.62672
##
```

TIM (Terrestrial Investigation Model)

By default, TIM estimates exposures for birds. The dermal exposure dose (body burden) is computed by combining the pesticide application rate (app_rate), surface area (sa_default_tim), fraction of surface area exposed (sa_tim_default_frac), body weight of the exposed organism (body_weight) and the dermal absorption factor (dermal_af).

$$TIM = \frac{AR * SA * SAF * DAF}{BW}$$

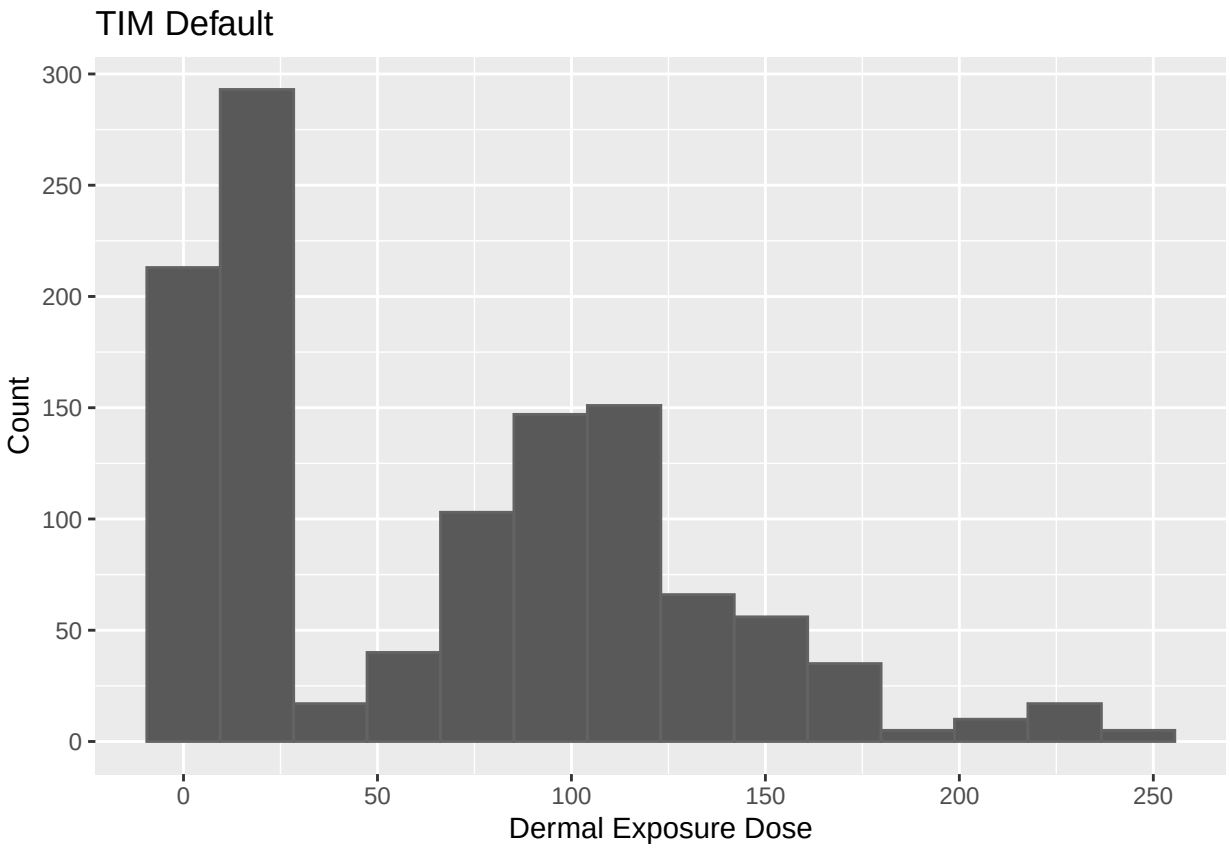
TIM Default Calculation

```
# Equation 1: TIM Default Values
app_rate <- combined_data$app_rate_g_cm2
body_weight <- combined_data$body_weight_g
conv_rate <- 1000000.0
dermal_af <- 1.0
hours_one <- 1.0
sa_default_tim <- 10 * (body_weight ^ 0.667)
sa_tim_default_frac <- 0.5
tissue_conc <- combined_data$tissue_conc_ugg
```

```
# Calculate Equation 1a: TIM Default
combined_data$tim_direct_default <- (app_rate * conv_rate * sa_default_tim * sa_tim_default_frac * dermal_af)/body_
tim_direct_default <- combined_data$tim_direct_default
```

```
# compute number of histogram bins
bw <- 2*IQR(tim_direct_default) / length(tim_direct_default)^(1/3)

# create frequency histogram
ggplot(combined_data, aes(x=tim_direct_default)) +
  geom_histogram(aes(y = ..count..), color = "#636363", binwidth = bw) +
  scale_x_continuous(name = "Dermal Exposure Dose", breaks = seq(0, 300, 50)) +
  scale_y_continuous(name = "Count", breaks = seq(0, 300, 50)) +
  ggtitle("TIM Default")
```



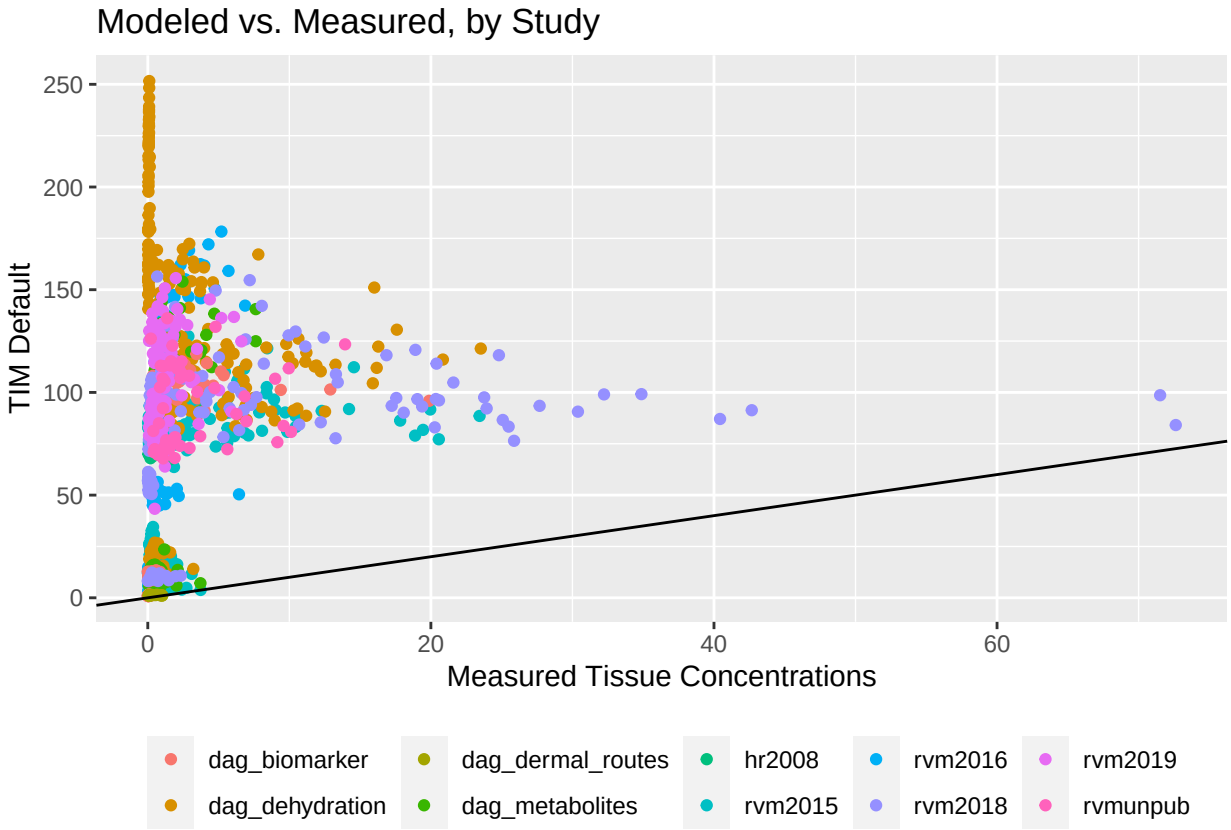
TIM Default Ratios

Here we compute and visualize the ratio of the TIM Default value to the measured tissue concentrations. A 1:1 ratio signifies perfect agreement between the modeled and measured values.

```
# Calculate Equation 1b: TIM Default Ratios
combined_data$tim_direct_default_ratios <- combined_data$tim_direct_default/tissue_conc
```

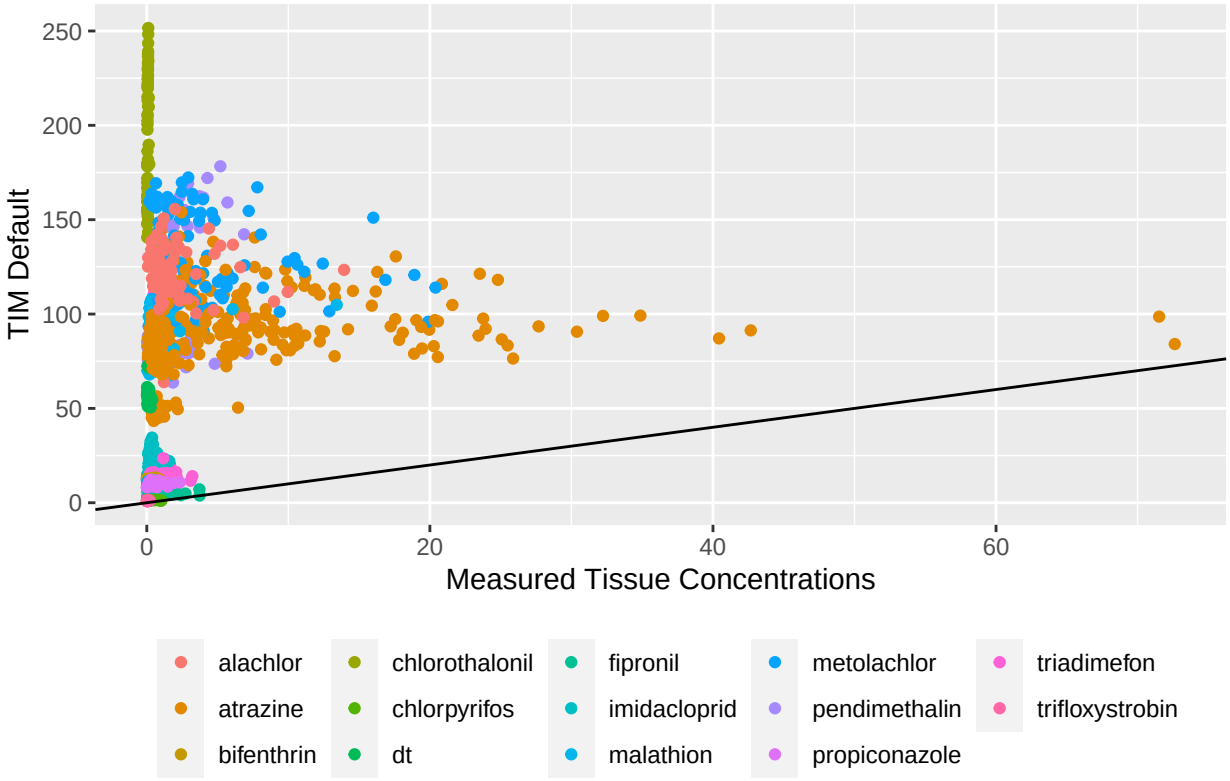
Scatterplots were created to visualize the model's performance, grouped by variables (study, chemical, application type), respectively. The trend line displays the 1:1 ratio.

```
# create scatterplot, categorized by study
ggplot(combined_data, aes(x = tissue_conc, y = tim_direct_default, color = source)) +
  geom_point() +
  labs(title = "Modeled vs. Measured, by Study", x = "Measured Tissue Concentrations", y = "TIM Default") +
  geom_abline(intercept = 0, slope = 1) +
  theme(legend.position = "bottom", legend.title = element_blank())
```

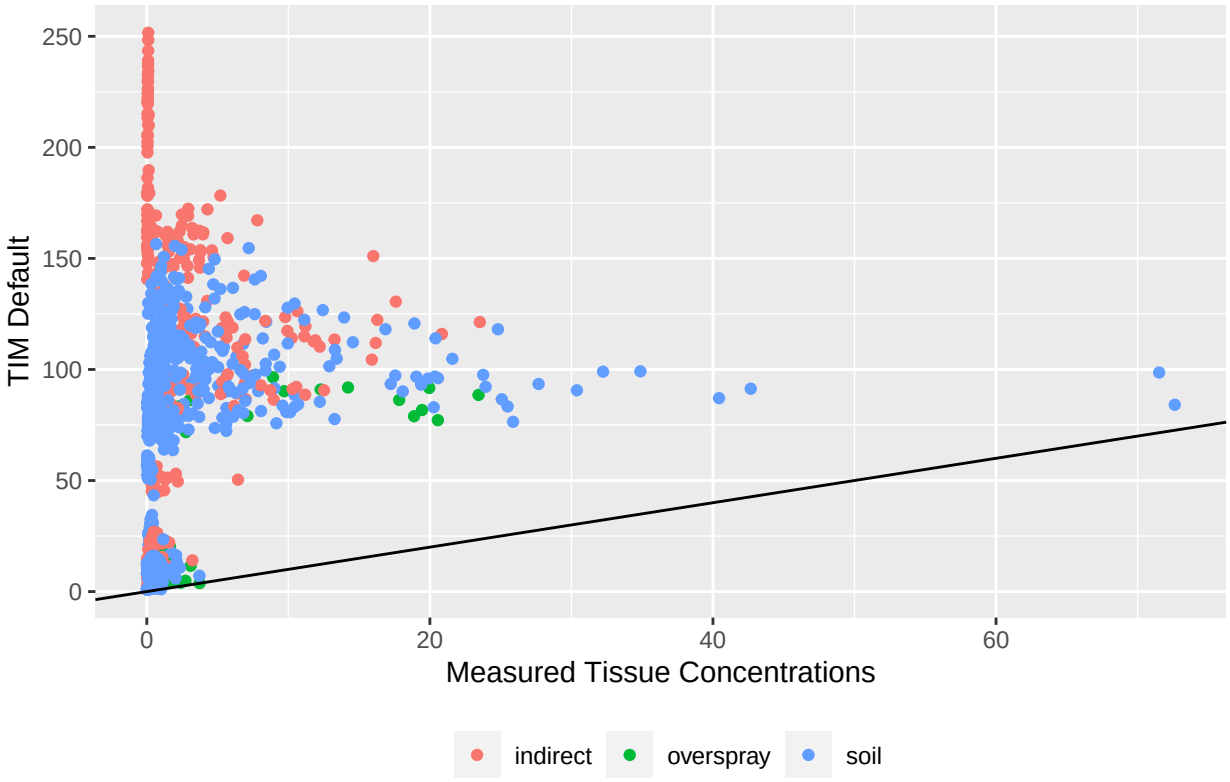



```
# create scatterplot, categorized by chemical
ggplot(combined_data, aes(x = tissue_conc, y = tim_direct_default, color = chemical)) +
  geom_point() +
  labs(title = "Modeled vs. Measured, by Chemical", x = "Measured Tissue Concentrations", y = "TIM Default") +
  geom_abline(intercept = 0, slope = 1) +
  theme(legend.position = "bottom", legend.title = element_blank())
```

Modeled vs. Measured, by Chemical



Modeled vs. Measured, by Application Type



TIM Amphibian Calculation

To gain a better understanding of pesticide risks on amphibians, the dermal exposure dose was also computed with a modified allometric relationship between body weight and surface area that is better suited for amphibians (Hutchison et al. 1968, EPA 2009).

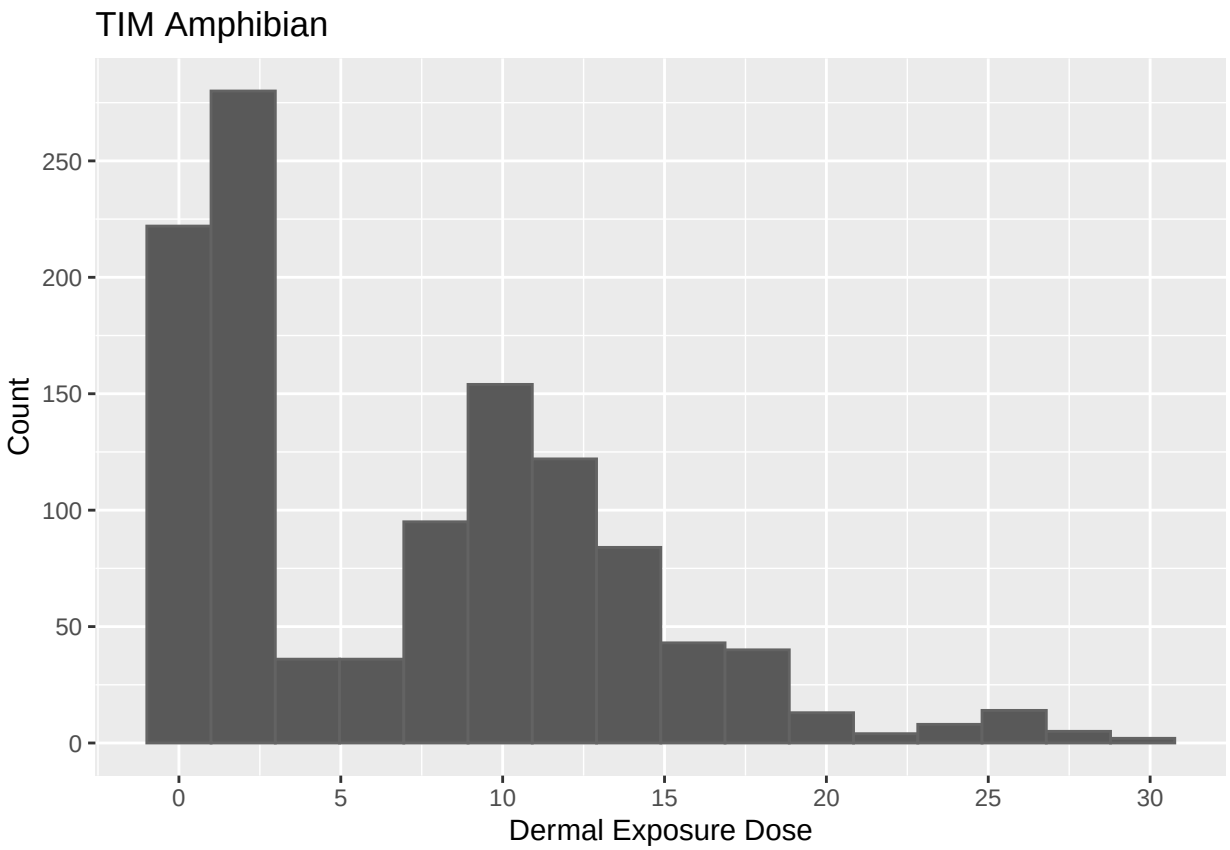
```
# Equation 2: TIM Amphibians value (Hutchinson body weight)
sa_amphib_hutchinson <- 1.131 * (body_weight ^ 0.579)
```

```
# Calculate Equation 2a: TIM Amphibian
combined_data$tim_direct_amphib <- (app_rate * conv_rate * sa_amphib_hutchinson * sa_tim_default_frac * dermal_af)

tim_direct_amphib <- combined_data$tim_direct_amphib
```

```
# compute number of bins
bw <- 2*IQR(tim_direct_amphib) / length(tim_direct_amphib)^(1/3)

# create frequency histogram
ggplot(combined_data, aes(x=tim_direct_amphib)) +
  geom_histogram(aes(y = ..count..), color = "#636363", binwidth = bw) +
  scale_x_continuous(name = "Dermal Exposure Dose", breaks = seq(0, 30, 5)) +
  scale_y_continuous(name = "Count", breaks = seq(0, 300, 50)) +
  ggtitle("TIM Amphibian")
```



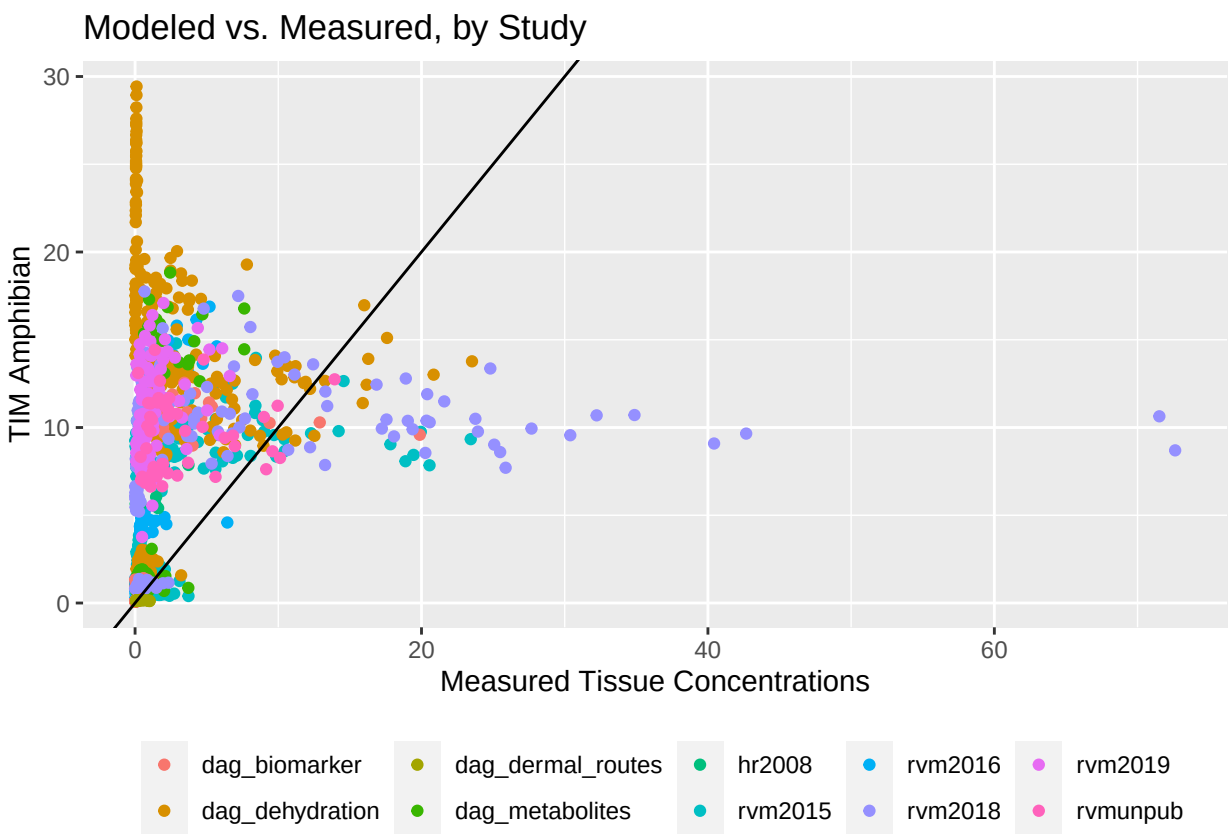
TIM Amphibian Ratios

Here we compute and visualize the ratio of the TIM Amphibian values to the measured tissue concentrations. A 1:1 ratio signifies perfect agreement between the modeled and measured values.

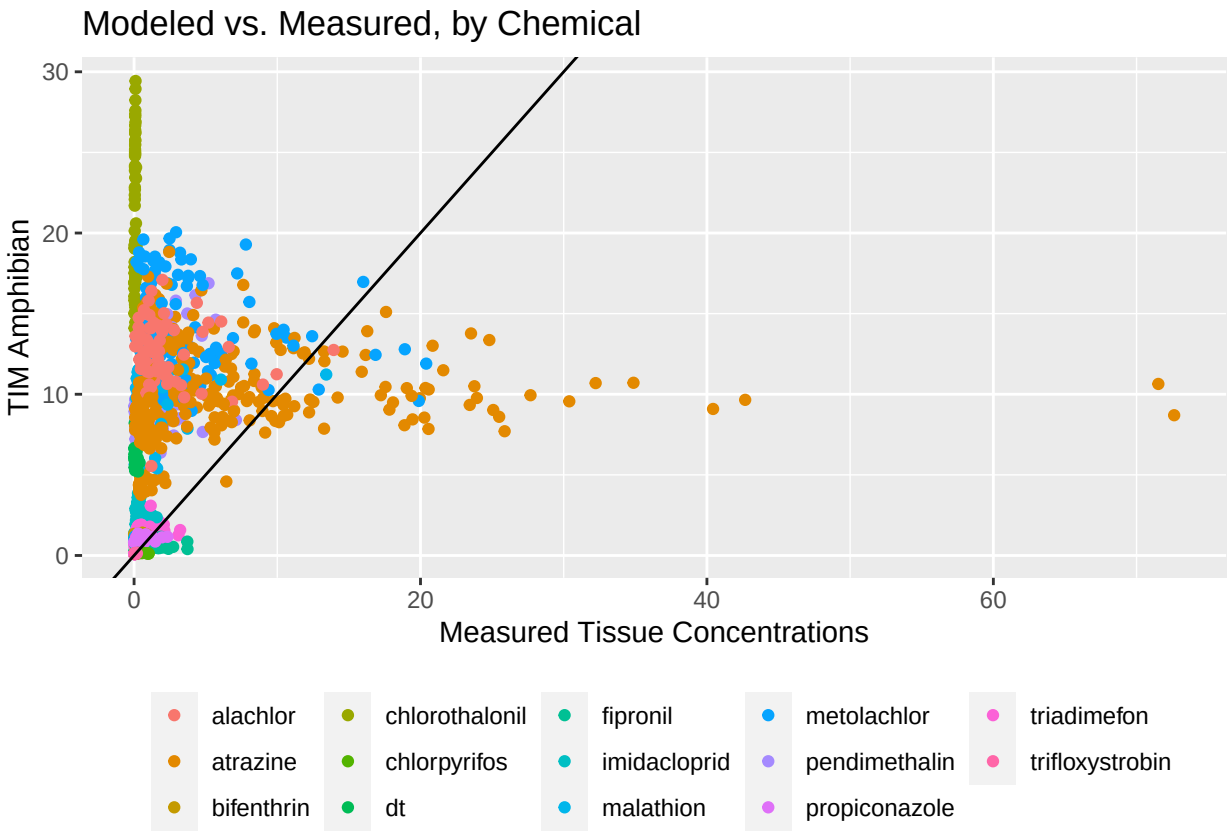
```r
# Calculate Equation 2b: TIM Amphibian Ratios
combined_data$tim_direct_amphib_ratios <- combined_data$tim_direct_amphib/tissue_conc
```

Scatterplots were created to visualize the model's performance, grouped by variables (study, chemical, application type), respectively. The trend line displays the 1:1 ratio.

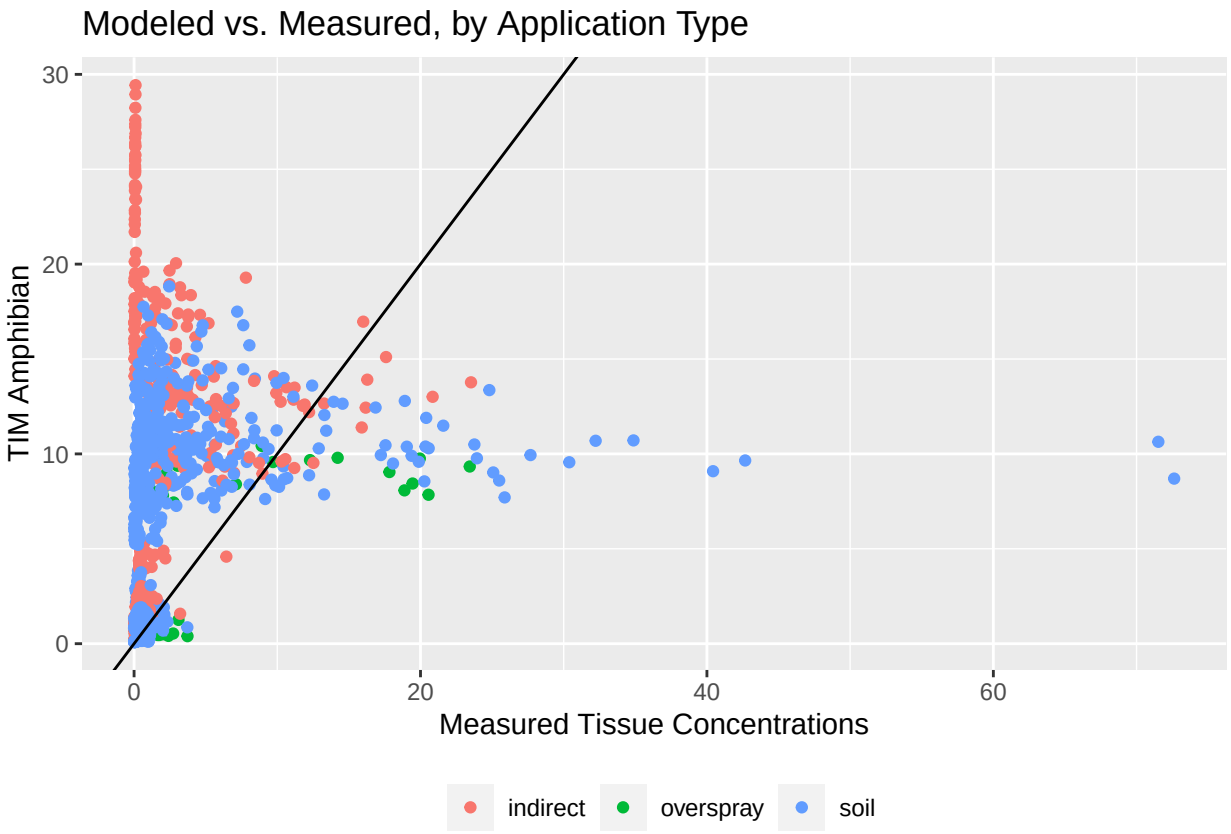
```r
# create scatterplot, categorized by study
ggplot(combined_data, aes(x = tissue_conc, y = tim_direct_amphib, color = source)) +
  geom_point() +
  labs(title = "Modeled vs. Measured, by Study", x = "Measured Tissue Concentrations", y = "TIM Amphibian") +
  geom_abline(intercept = 0, slope = 1) +
  theme(legend.position = "bottom", legend.title = element_blank())
```




```r
# create scatterplot, categorized by chemical
ggplot(combined_data, aes(x = tissue_conc, y = tim_direct_amphib, color = chemical)) +
  geom_point() +
  labs(title = "Modeled vs. Measured, by Chemical", x = "Measured Tissue Concentrations", y = "TIM Amphibian") +
  geom_abline(intercept = 0, slope = 1) +
  theme(legend.position = "bottom", legend.title = element_blank())
```




```r
# create scatterplot, categorized by application type
ggplot(combined_data, aes(x = tissue_conc, y = tim_direct_amphib, color = application)) +
  geom_point() +
  labs(title = "Modeled vs. Measured, by Application Type", x = "Measured Tissue Concentrations", y = "TIM Amphibia
  geom_abline(intercept = 0, slope = 1) +
  theme(legend.position = "bottom", legend.title = element_blank())
```



Assessing Model Performance

The TIM Default and TIM Amphibian models, described and shown above, were compared to one another and to the post-exposure amphibian body burdens found in the collated data set in order to assess model performance.

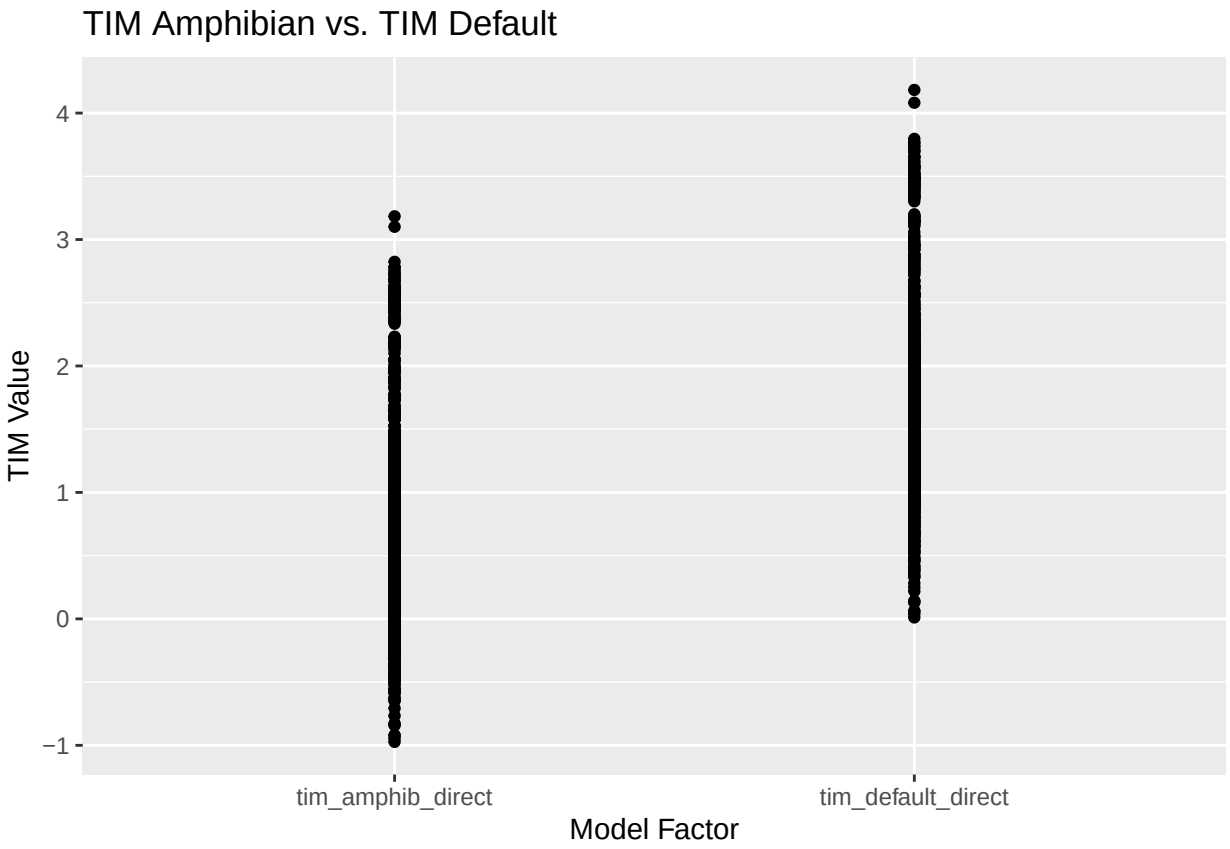
```
# Create df of ratios
model_log_ratios <- as.vector(c(log10(combined_data$tim_direct_default_ratios),
                      y=log10(combined_data$tim_direct_amphib_ratios)))
model_factor <- as.vector(c(rep('tim_default_direct',n),rep('tim_amphib_direct',n)))
df.ratios <- data.frame(x=model_factor, y=model_log_ratios)

# Create df of tissue_conc and TIM Amphibian
df.tim_amphib <- data.frame(x=tissue_conc, y=tim_direct_amphib)

# Create df of tissue_conc and TIM Default
df.tim_default <- data.frame(x=tissue_conc, y=tim_direct_default)
```

Comparing Models

```
# compare default and amphib ratios
ggplot(df.ratios, aes(x=x, y=y)) +
  geom_point() +
  labs(title="TIM Amphibian vs. TIM Default",  x="Model Factor", y = "TIM Value")
```



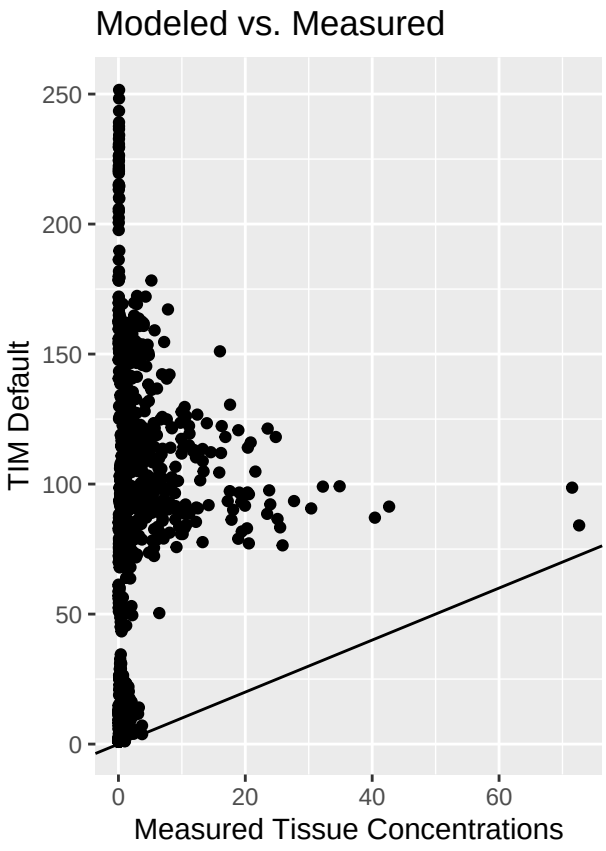
Plotting Modeled vs. Measured

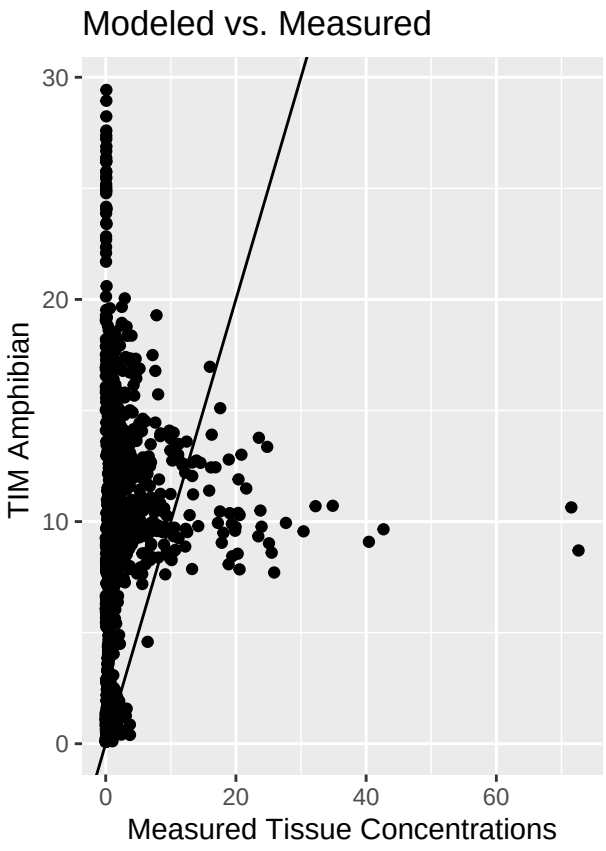
Again, scatterplots comparing the modeled vs. measured values are shown, this time setting models side-by-side. A model along the 1:1 line would display perfect agreement with the measured values.

```
# scatterplot of TIM default values
plot_default <- ggplot(df.tim_default, aes(x=x, y=y)) +
  geom_point() +
  labs(title="Modeled vs. Measured",  x="Measured Tissue Concentrations", y = "TIM Default") +
  geom_abline(intercept = 0, slope = 1)

# scatterplot of TIM amphib values
plot_amphib <- ggplot(df.tim_amphib, aes(x=x, y=y)) +
  geom_point() +
  labs(title="Modeled vs. Measured",  x="Measured Tissue Concentrations", y = "TIM Amphibian ") +
  geom_abline(intercept = 0, slope = 1)

# arrange 1row*2col grid for display
grid.arrange(plot_default, plot_amphib, ncol = 2)
```

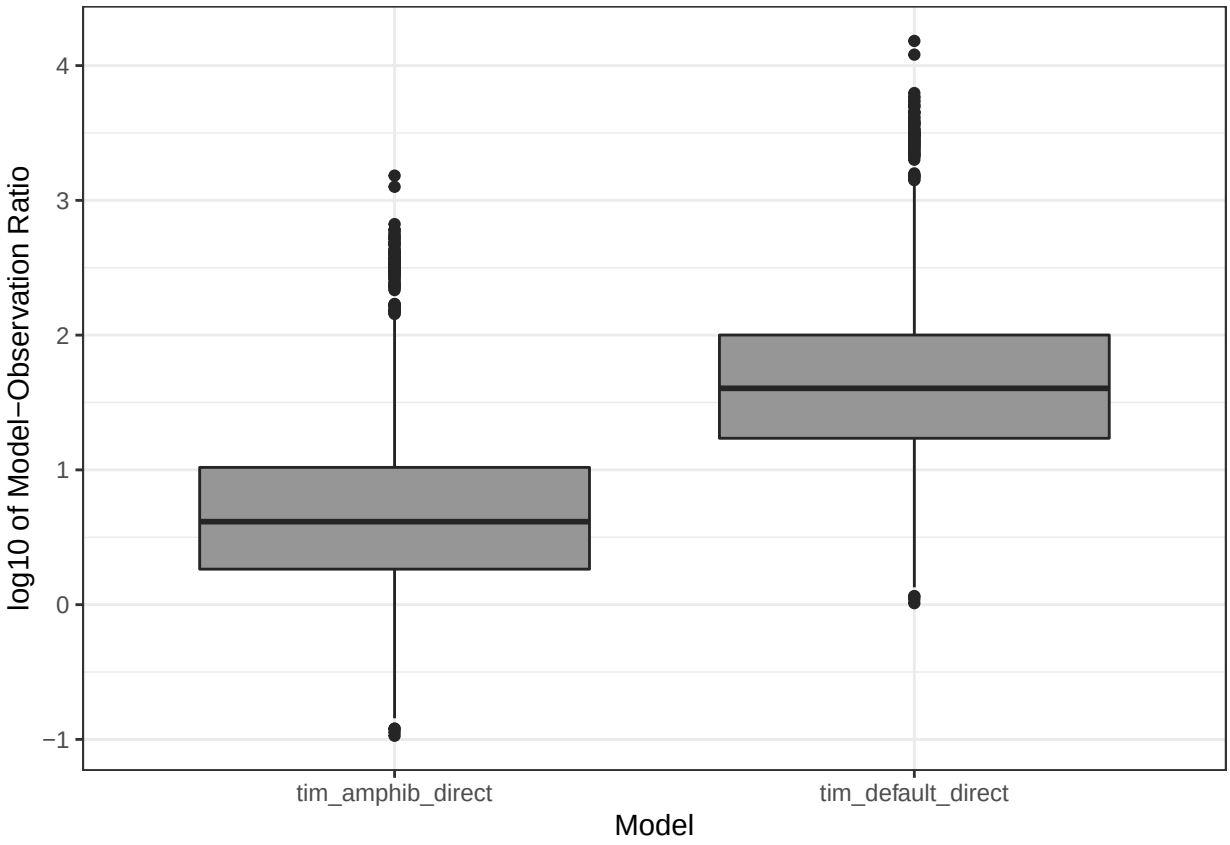




Plotting log10 Ratio of Modeled Burdens to Observed Burdens

Box plots were utilized to further assess model performance. Well-performing models would minimize error, and therefore would have a higher percentage of ratios near zero (log10 of unity).

```
fill <- "#969696"
line <- "#252525"
ggplot(df.ratios, aes(x=x,y=y)) +
    geom_boxplot(fill = fill, colour = line) +
    scale_x_discrete(name = "Model") +
    scale_y_continuous(name = "log10 of Model-Observation Ratio") +
    theme_bw()
```




```
model_boxplot_filename <- paste(amphibdir_graphics,"model_boxplot.png",sep='')
ggsave(model_boxplot_filename, device="png", width=4, height=4)
```

False Negative Rates

Summary statistics for false negative rates were also computed for both models. False negative rates are defined as an outcome where the modeled value is less than the measured value. This rate is based on a Type II Error in statistical hypothesis testing. False negative rates assess how often a model under-estimates the measurement of interest.

TIM Default As displayed below, there are zero false negative rates for TIM Default. But, 138/1158 (11.9%) of the measurements are within an order of magnitude of the modeled value.

```
# number of samples
n
```

```
## [1] 1158
```

```
# False Negative Rate: TIM Default
sum(tissue_conc > tim_direct_default)
```

```
## [1] 0
```

```
sum(tissue_conc > tim_direct_default)/n
```

```
## [1] 0
```

```
# False Negative Rate (x10): TIM Default
sum(tissue_conc*10 > tim_direct_default)
```

```
## [1] 138
```

```
sum(tissue_conc*10 > tim_direct_default)/n
```

```
## [1] 0.119171
```

TIM Amphibian There are 131/1158 (11.3%) false negative screen results for TIM Amphibian, and 786/1014 (77.5%) measurements within an order of magnitude of the modeled value. TIM Amphibian uses a less conservative but more representative surface area calculation for amphibians.

```
#number of samples
n
```

```
## [1] 1158
```

```
# False Negative Rate: TIM Amphibian
sum(tissue_conc > tim_direct_amphib)
```

```
## [1] 131
```

```
sum(tissue_conc > tim_direct_amphib)/n
```

```
## [1] 0.1131261
```

```
# False Negative Rate (x10): TIM Amphibian
sum(tissue_conc*10 > tim_direct_amphib)
```

```
## [1] 863
```

```
sum(tissue_conc*10 > tim_direct_amphib)/n
```

```
## [1] 0.7452504
```

By Study False Negative Rates were also computed indiviually by study source. The table below displays the false negative rates and the measurements within an order of magnitude of the modeled value (displayed as %).

```
## Warning: package 'knitr' was built under R version 4.0.3
```

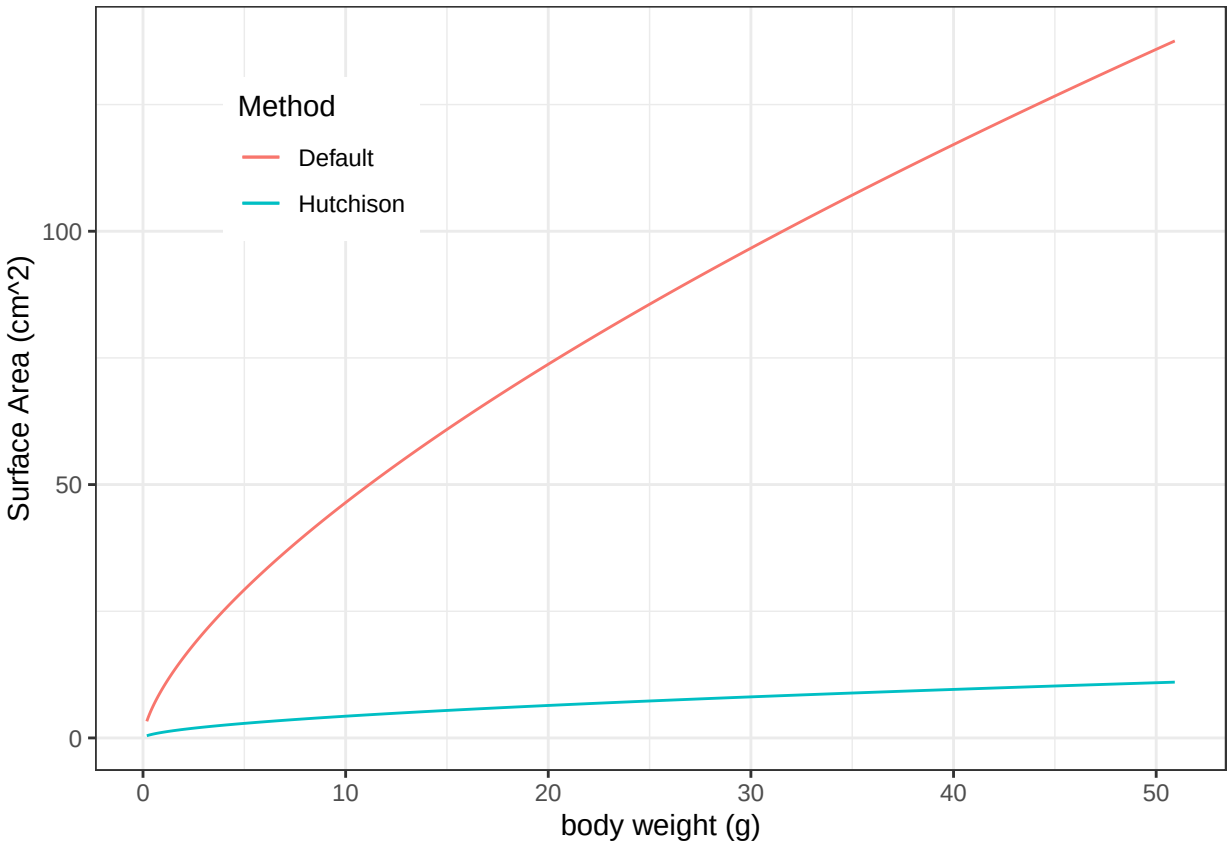
```
## Warning: package 'kableExtra' was built under R version 4.0.5
```

		TIM Default		TIM Amphibian	
	sample_size	fnr	fnr_10	fnr	fnr_10
Van Meter 2014/15 total	196	0	25.00	22.45	78.06
Van Meter 2014/15 soil	151	0	17.22	13.91	72.85
Van Meter 2014/15 overspray	45	0	51.11	51.11	95.56
Van Meter 2016	96	0	1.04	2.08	76.04
Van Meter 2018	0	NaN	NaN	NaN	NaN
Henson-Ramsey 2008	9	0	0.00	0.00	55.56
Glinski 2018b (Metabolites)	60	0	18.33	16.67	88.33
Glinski 2018a (Dehydration)	300	0	6.00	5.00	69.33
Glinski 2019 (Biomarkers)	192	0	3.65	4.17	94.27
Glinski 2020 (Dermal)	24	0	50.00	50.00	100.00

Comparing Surface Area Calculations

While TIM Amphibian had a higher percentage of false negative screening values than TIM Default, the difference in screening values can be attributed to the surface area calculation. Amphibian surface area, as compared to the TIM Default surface area based on birds, is extrapolated from measured body weights. The comparison between surface area calculations is depicted below.

```
min_bw <- min(combined_data$body_weight_g, na.rm=T)
max_bw <- max(combined_data$body_weight_g, na.rm=T)
range_bw <- seq(min_bw, max_bw, by = 0.01)
default_sa <- 10*range_bw^0.667
hutchison_sa <- 1.131 * range_bw^0.579
#max(default_sa/hutchison_sa)
sa_data <- data.frame(range_bw, default_sa, hutchison_sa)
ggplot(sa_data, aes(range_bw)) +
  labs(x = "body weight (g)", y = "Surface Area (cm^2)",color = "Method") +
  geom_line(aes(y = default_sa, colour = "Default")) +
  geom_line(aes(y = hutchison_sa, colour = "Hutchison")) +
  theme_bw() +
  theme(legend.position =c(0.2,0.8))
```




```
sa_comparison_filename <- paste(amphibdir_graphics,"sa_comparison.png",sep='')
ggsave(sa_comparison_filename, device="png", width=4, height=4)
```

Session Information

```
## R version 4.0.2 (2020-06-22)
## Platform: x86_64-w64-mingw32/x64 (64-bit)
## Running under: Windows 10 x64 (build 18363)
##
## Matrix products: default
##
## locale:
## [1] LC_COLLATE=English_United States.1252
## [2] LC_CTYPE=English_United States.1252
## [3] LC_MONETARY=English_United States.1252
## [4] LC_NUMERIC=C
## [5] LC_TIME=English_United States.1252
##
## attached base packages:
## [1] stats     graphics  grDevices utils     datasets  methods   base
##
## other attached packages:
## [1] kableExtra_1.3.4 knitr_1.31       tinytex_0.32     reshape2_1.4.4
## [5] gridExtra_2.3    ggplot2_3.3.3
##
## loaded via a namespace (and not attached):
##  [1] Rcpp_1.0.6        highr_0.8         pillar_1.6.0      compiler_4.0.2
##  [5] plyr_1.8.6        tools_4.0.2       digest_0.6.27     viridisLite_0.3.0
```

```
##  [9] evaluate_0.14     lifecycle_1.0.0   tibble_3.1.1      gtable_0.3.0
## [13] pkgconfig_2.0.3   rlang_0.4.10      rstudioapi_0.13   DBI_1.1.1
## [17] yaml_2.2.1        xfun_0.24         xml2_1.3.2        httr_1.4.2
## [21] withr_2.4.1       stringr_1.4.0     dplyr_1.0.5       systemfonts_1.0.2
## [25] generics_0.1.0    vctrs_0.3.7       webshot_0.5.2     grid_4.0.2
## [29] tidyselect_1.1.0  svglite_2.0.0     glue_1.4.1        R6_2.5.0
## [33] fansi_0.4.2       rmarkdown_2.6     farver_2.0.3      purrr_0.3.4
## [37] magrittr_2.0.1    scales_1.1.1      ellipsis_0.3.1    htmltools_0.5.1.1
## [41] rvest_0.3.6       assertthat_0.2.1  colorspace_1.4-1  labeling_0.3
## [45] utf8_1.2.1        stringi_1.5.3     munsell_0.5.0     crayon_1.4.1
```