Effects of urban green infrastructure on supply provision of ecosystem service: Temperature reduction

Irina Lerner

30/01/24

Contents

Aim and scope

CLEAR AIM: TO SHOW THE EFFECT OF VEGETATION CONFIGURATION IN MEAN TEMPERATURE.

Here we show results, flow and maybe missing steps

Data

DATA NEEDS URGENT CARE: - SEE IF ITS BETTER USING VI PROPORTIONS WITH OR WITHOUT NA. PROBLY WITH NA IF WE ARE CONSIDERING OTHER VARIABLES - PROP_WATER AND PROP_GRASS THAT ARE NOT CORRECTLY CALCULATED FROM THE MAPS

Exploratory Analysis

Response

We begin exploring the response data.

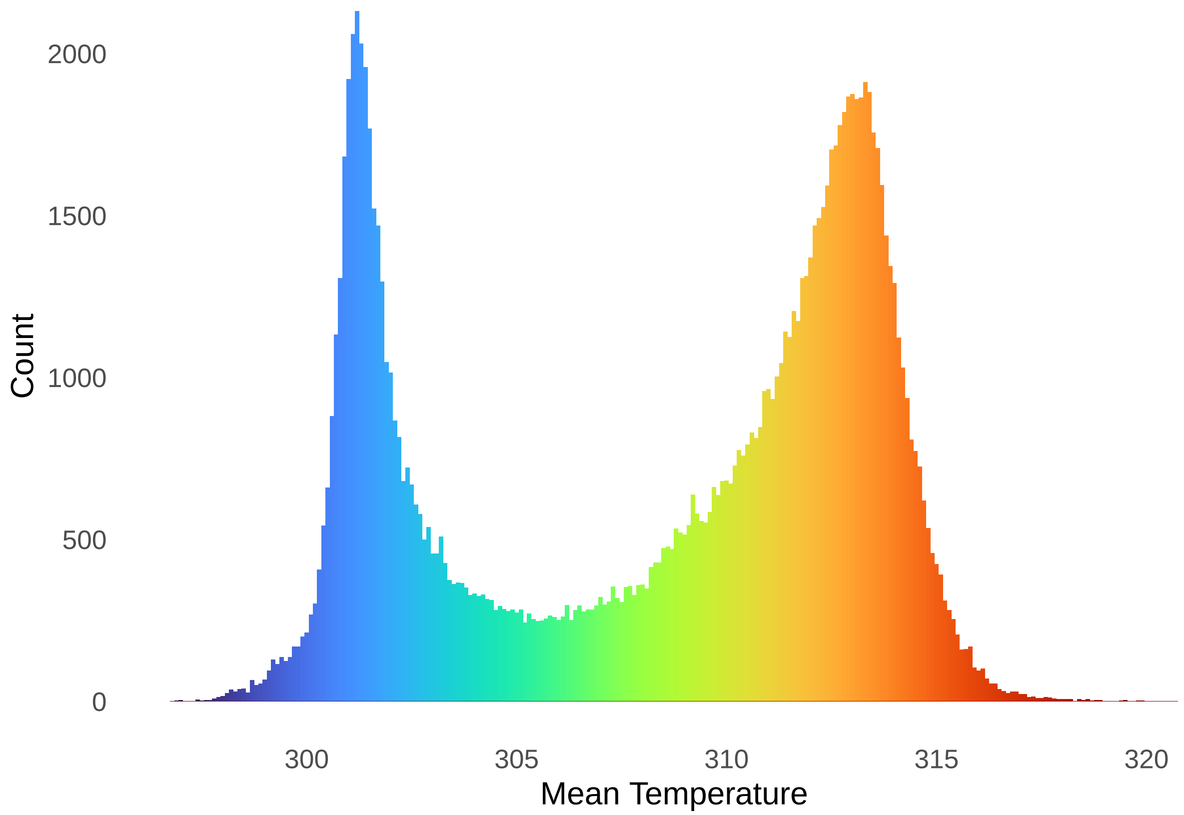
```
analyse_model(data, lm(T_mean ~ 1, data))
```

```
## [1] "Residuals not normally distributed:"
## [1] -0.2141189
##
##  Moran I test under randomisation
##
## data:  residuals(model)
## weights: listw
##
## Moran I statistic standard deviate = 736.58, p-value < 2.2e-16
## alternative hypothesis: greater
## sample estimates:
## Moran I statistic       Expectation          Variance
##      9.653109e-01     -6.748504e-06      1.717531e-06
```

```
## [1] 898590.2
```

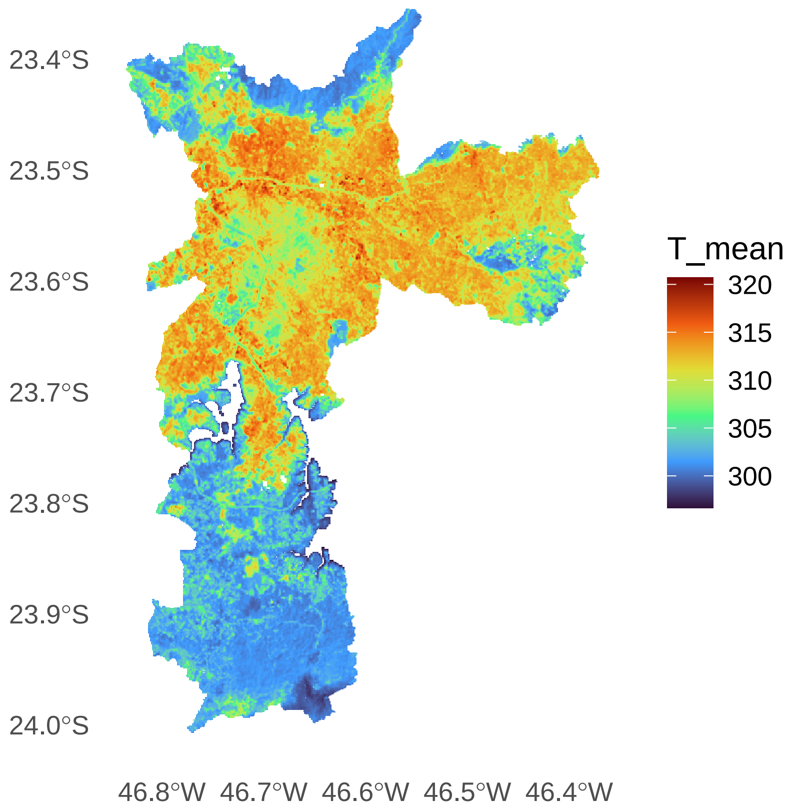
```
ggplot(cdata, aes(x = T_mean)) +
      geom_histogram(aes(fill = as.factor(cut_width(T_mean, 0.1))), binwidth = 0.1, color = NA) +
      scale_fill_viridis_d(option = "turbo", guide = FALSE) +
      labs(title = "Distribution of T_mean by Bin Order", x = "Mean Temperature", y = "Count") +
      custom_theme
```

```
## Warning: The `guide` argument in `scale_*()` cannot be `FALSE`. This was deprecated in
## ggplot2 3.3.4.
## i Please use "none" instead.
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was
## generated.
```



From the distribution of DT in this grid comparison, the values are not normally distributed, instead, they have two peaks.

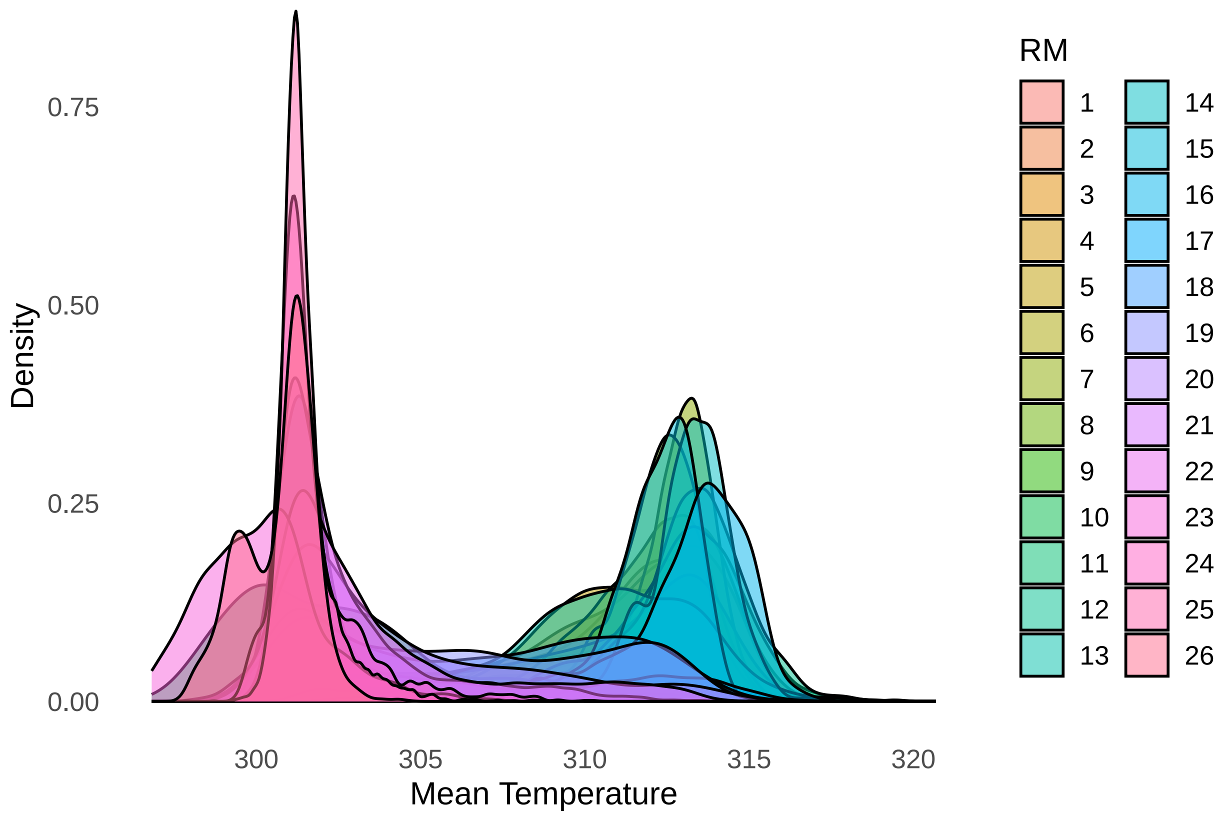
```
ggplot(data) +
  geom_sf(aes(fill = T_mean), color = NA) +  # Correct usage of color setting
  scale_fill_viridis_c(option = "turbo") +
  custom_theme
```



Predictors

This two peaks can be seen when we take the distribution of each micro climatic region

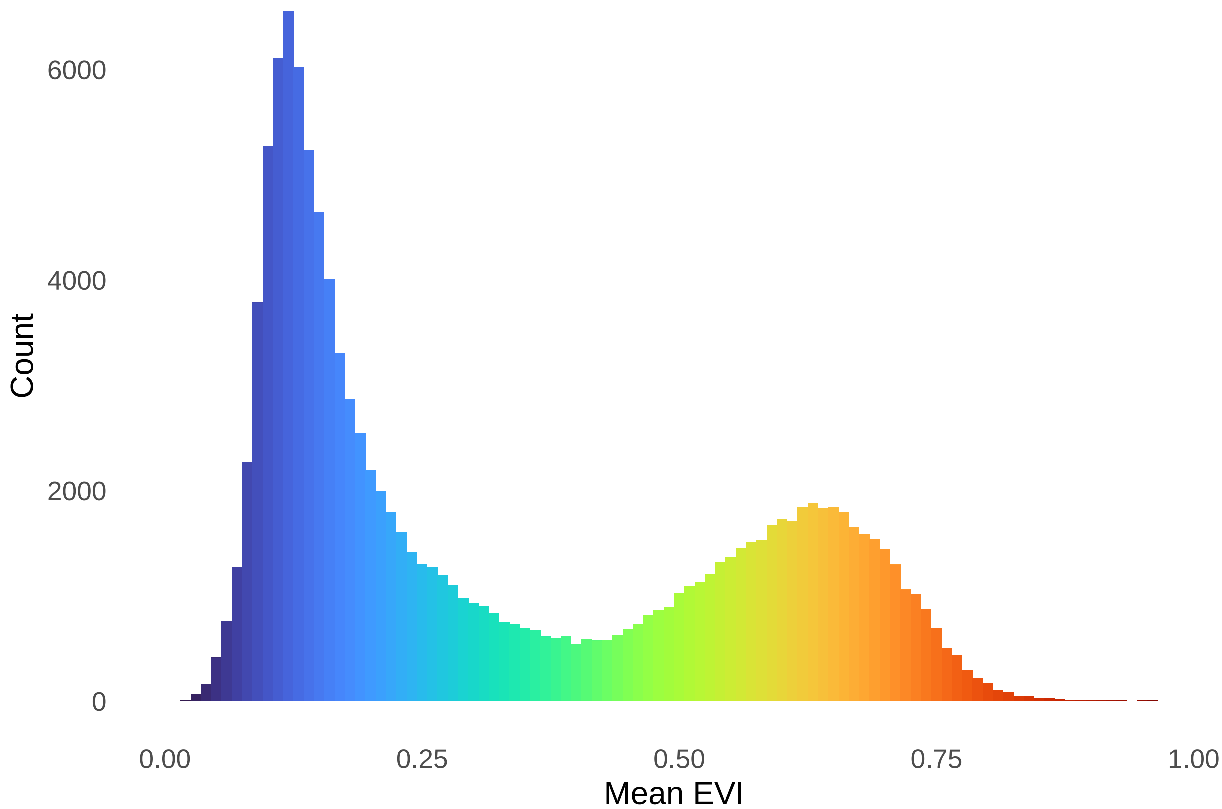
```
ggplot(cdata, aes(x = T_mean, fill = factor(RM))) +
  geom_density(alpha = 0.5) +
  labs(x = "Mean Temperature", y = "Density", fill = "RM") +
  custom_theme
```



we can either normalize by micro climatic region or consider it as a random effect in the distribution parameters.

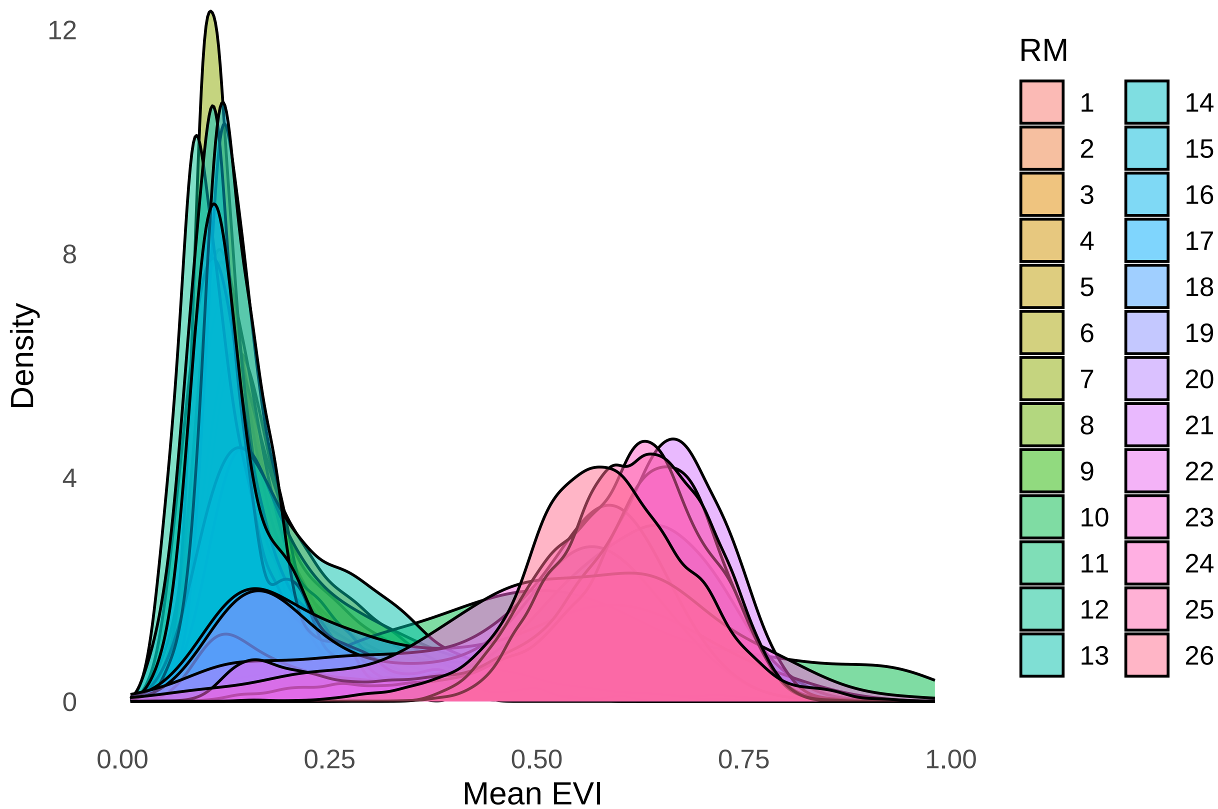
But looking at the distribution of São Paulo mean EVI values in the 100m² grid, we see again the peaks,

```
ggplot(cdata, aes(x = EVI_mean)) +
      geom_histogram(aes(fill = as.factor(cut_width(EVI_mean, 0.01))), binwidth = 0.01, color = NA) +
      scale_fill_viridis_d(option = "turbo", guide = FALSE) +
      labs(x = "Mean EVI", y = "Count") +
      custom_theme
```



Plotting the histogram divided by micro climatic region gives us the peaks again, indicating similar pattern found in the last image.

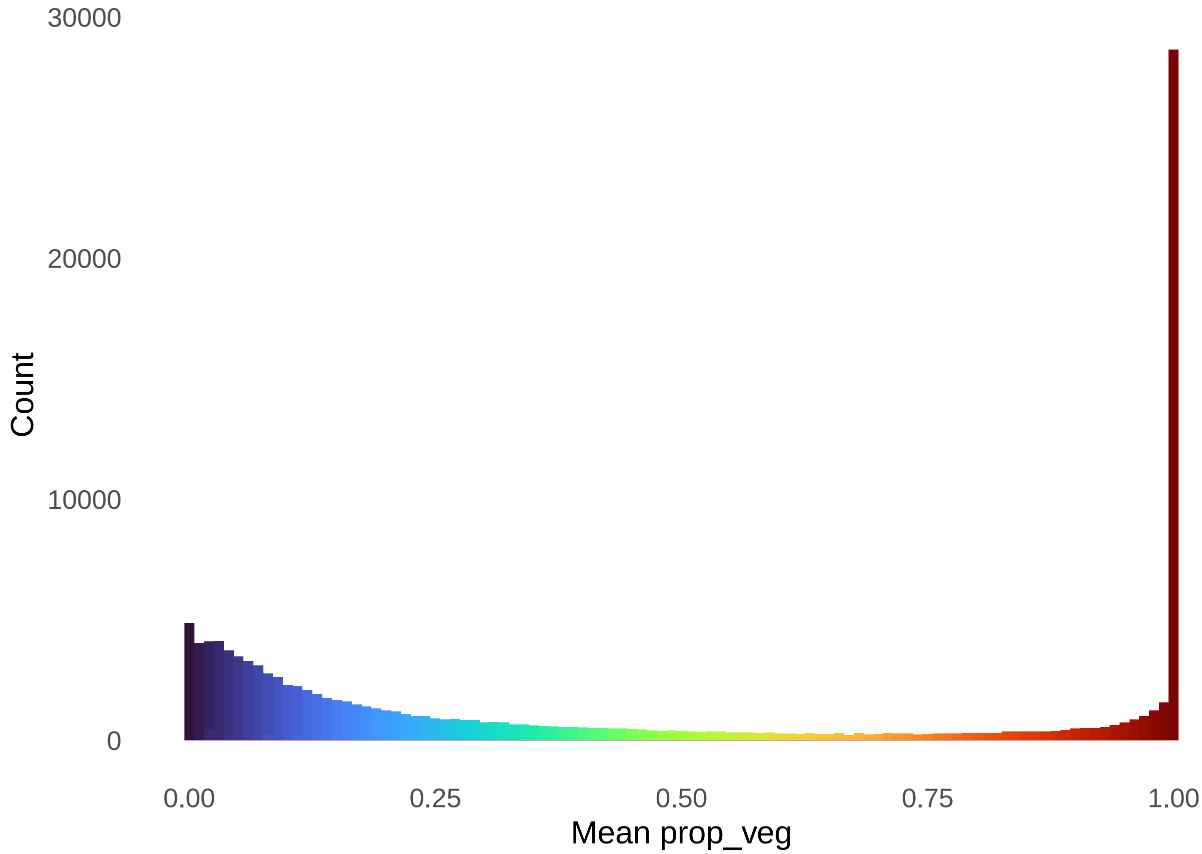
```
ggplot(cdata, aes(x = EVI_mean, fill = factor(RM))) +
  geom_density(alpha = 0.5) +
  labs(title = "Density Plot of mean EVI by Microclimatic region", x = "Mean EVI", y = "Density", fill =
  custom_theme
```



This implies that micro climatic regions are way different in vegetation and that might be what is responsible for this situation. But we can visually conclude that the number of places with low temperature peaks higher than the number of places with high EVI.

When we look at proportion of vegetation distribution we see what might lie behind this effect. While the high temperature and low vegetation peaks kind of match, the high EVI doesn't correspond to the peak, but wee see that São Paulo has more fully vegetated patches than anything else.

```
ggplot(cdata, aes(x = prop_veg)) +
      geom_histogram(aes(fill = as.factor(cut_width(prop_veg, 0.01))), binwidth = 0.01, color = NA) +
      scale_fill_viridis_d(option = "turbo", guide = FALSE) +
      labs(title = "Distribution of prop_veg by Bin Order", x = "Mean prop_veg", y = "Count") +
      custom_theme
```

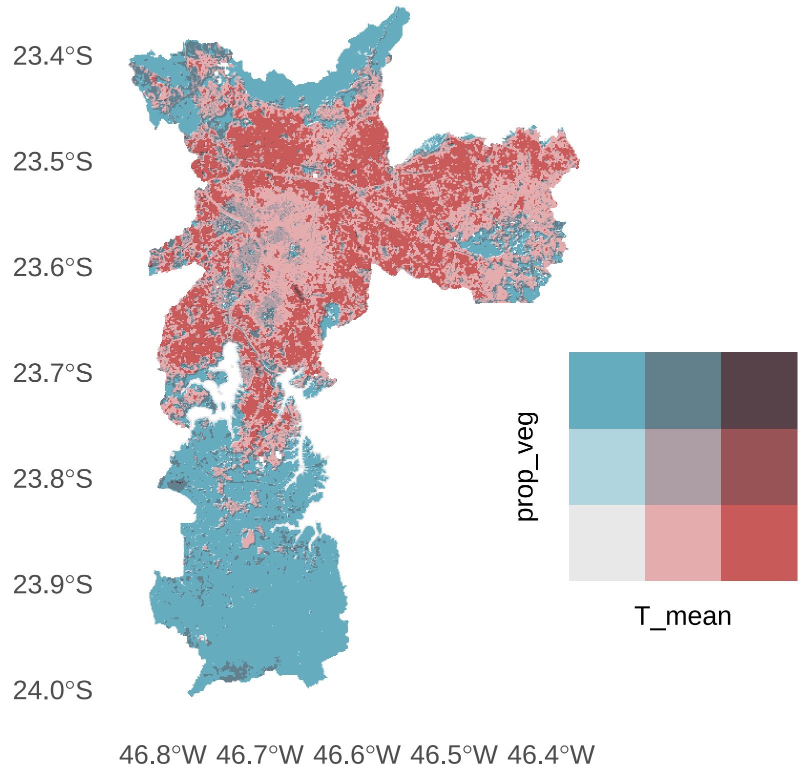


Some hypothesis can be drawn from that. In particular, that the effect of vegetation in temperature is not linear, but increases with the increase in vegetation. Also, that we might be seeing an spatial effect of vegetation, where it not only reduces local temperature but can cool down neighboring regions. Lets explore those hypothesis.

Let's see the maps overlaying

```r
plot_bi_map(data, "T_mean", "prop_veg")
```

```
##
##   1-1   1-2   1-3   2-1   2-2   2-3   3-1   3-2   3-3
##  2487  2455 46675 35988 13712 12059 32919  1686   201
```

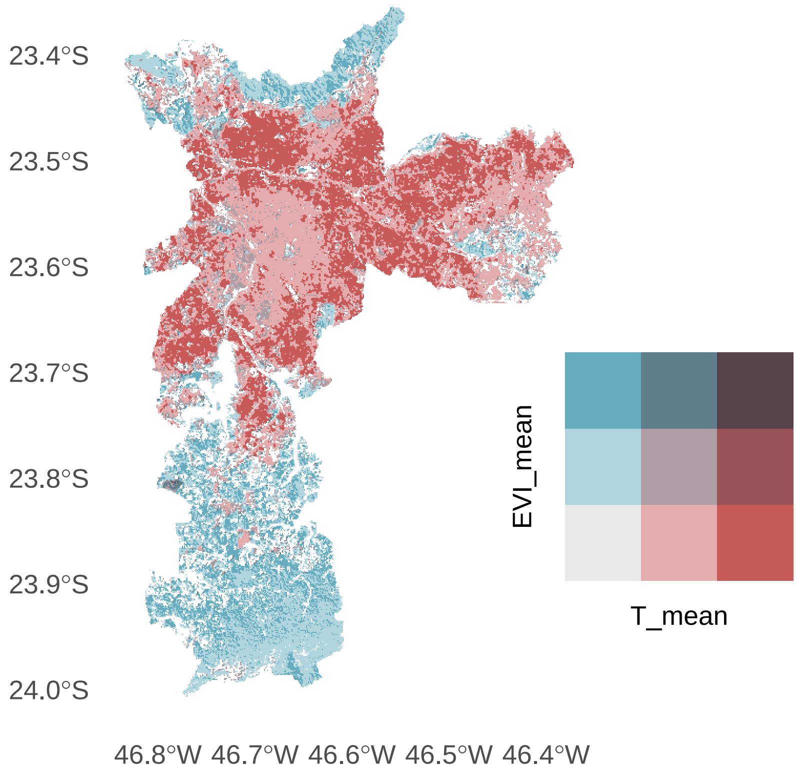


Combinations of low proportion of vegetation and low mean temperature seem to be around water, forest borders, and in a specific region in the center of the city. In the reserves, we see low temperature and high vegetation. In the urban part of the city we see low vegetation with high temperature as expected.

We can also check the biplots for other variables

```
plot_bi_map(cdata, "T_mean", "EVI_mean")
```

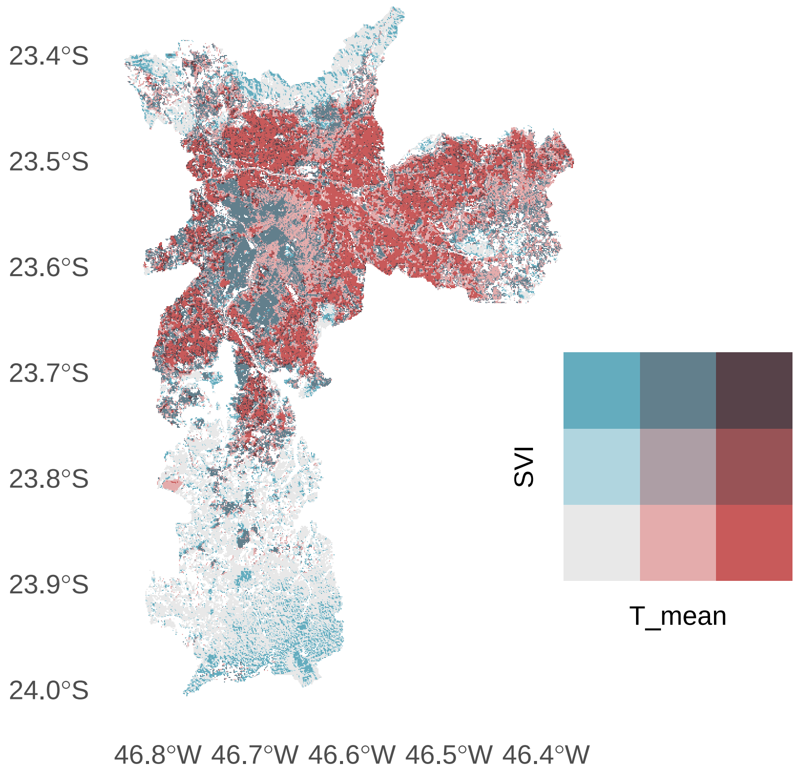
```
##
##    1-1   1-2   1-3   2-1   2-2   2-3   3-1   3-2   3-3
##    883 25412 13569 40217  9598  1076 32299   428    18
```



Lets check for our response variables

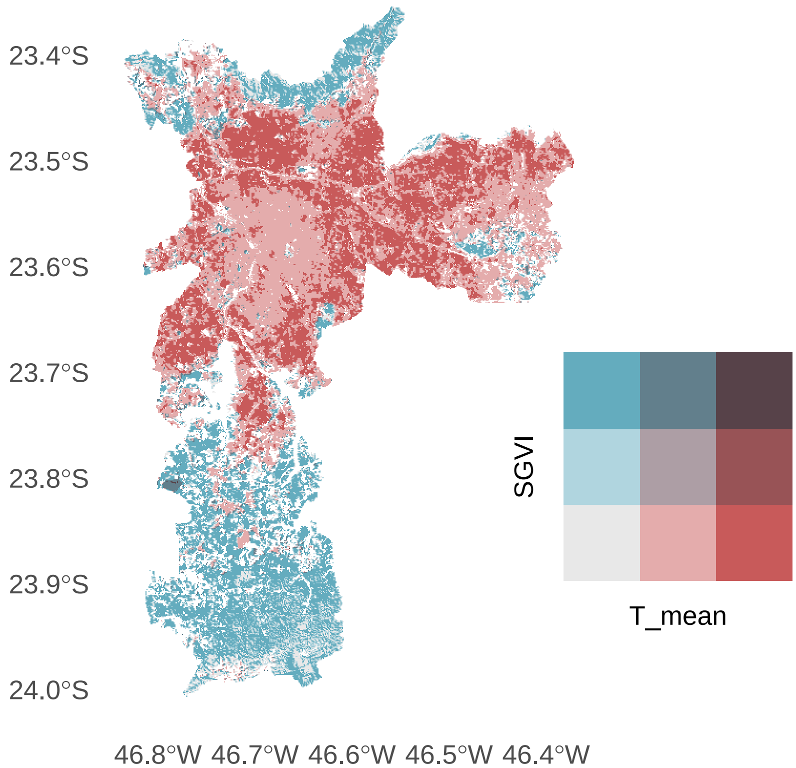
```
plot_bi_map(cdata, "T_mean", "SVI")
```

```
##
##   1-1   1-2   1-3   2-1   2-2   2-3   3-1   3-2   3-3
## 27288  6417  6159 19564 13044 18283 22131  6723  3891
```



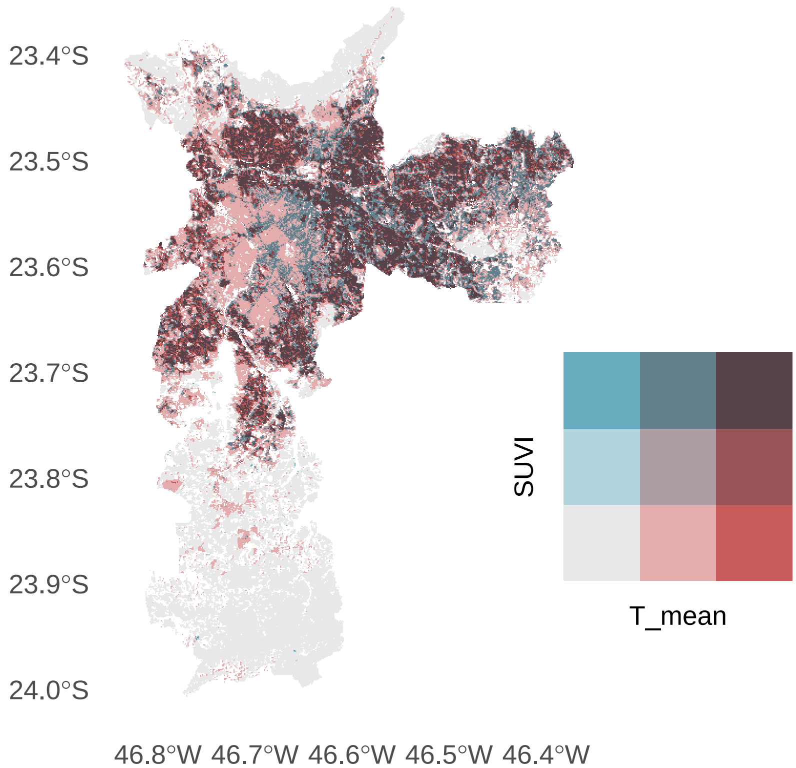

```
plot_bi_map(cdata, "T_mean", "SGVI")
```

```
##
##   1-1   1-2   1-3   2-1   2-2   2-3   3-1   3-2   3-3
##  7521  6462 25881 46602  2010  2279 32651    64    30
```



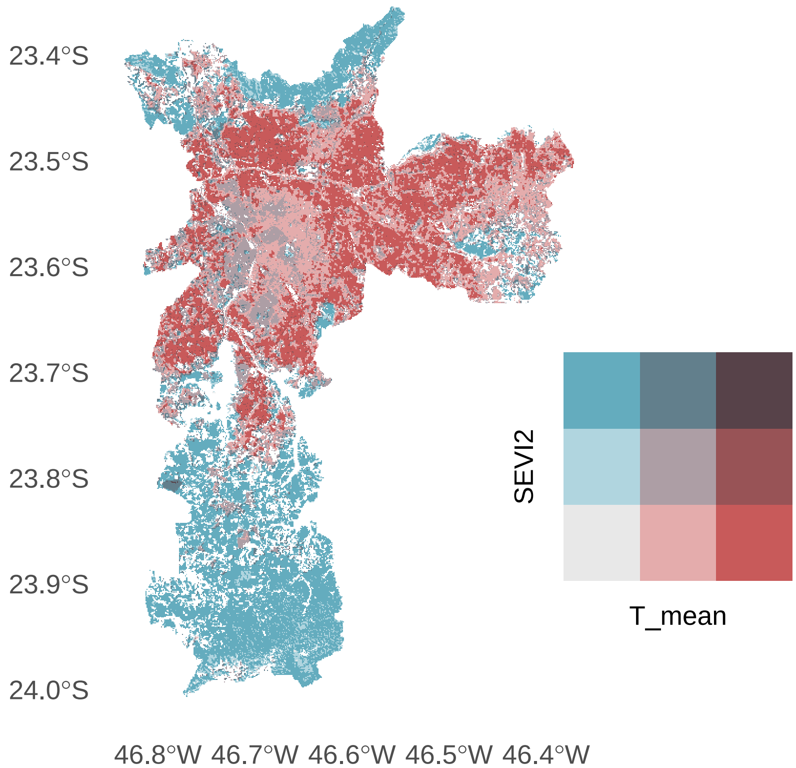

```r
plot_bi_map(cdata, "T_mean", "SUVI")
```

```
##
##   1-1   1-2   1-3   2-1   2-2   2-3   3-1   3-2   3-3
## 39630   153    81 26020  9853 15018  4870  6736 21139
```




```
plot_bi_map(cdata, "T_mean", "SEVI2")
```

```
##
##   1-1   1-2   1-3   2-1   2-2   2-3   3-1   3-2   3-3
##   267  6129 33468 26938 19253  4700 29045  3596   104
```



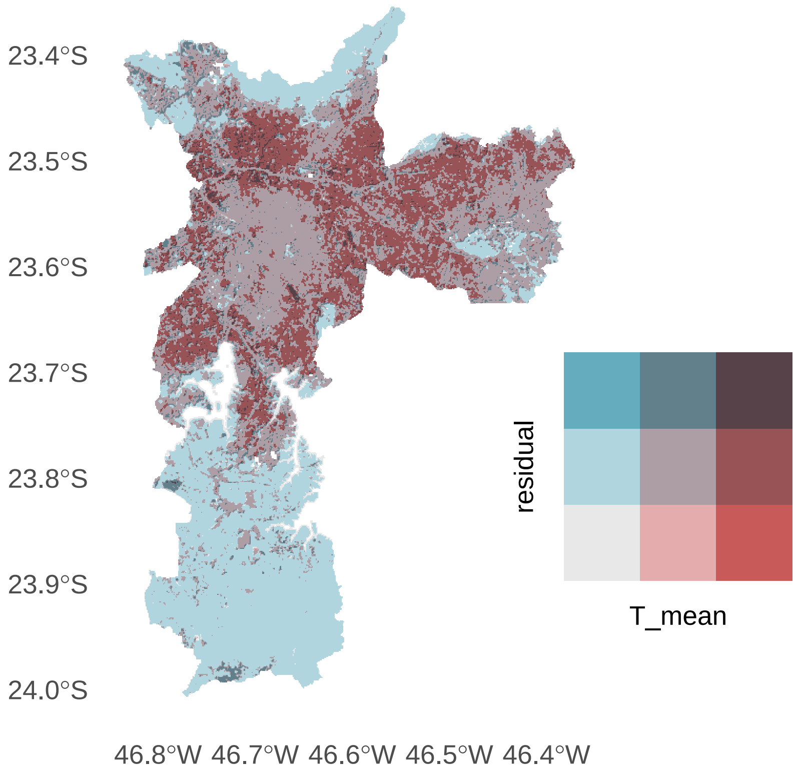
Those maps show us important features about the homogenity of the process that is happening in the city. Let's check the residuals of a linear model:

```r
linear <- lm(T_mean ~ prop_veg, data)

data$residual <- resid(linear)

plot_bi_map(data, "T_mean", "residual")
```

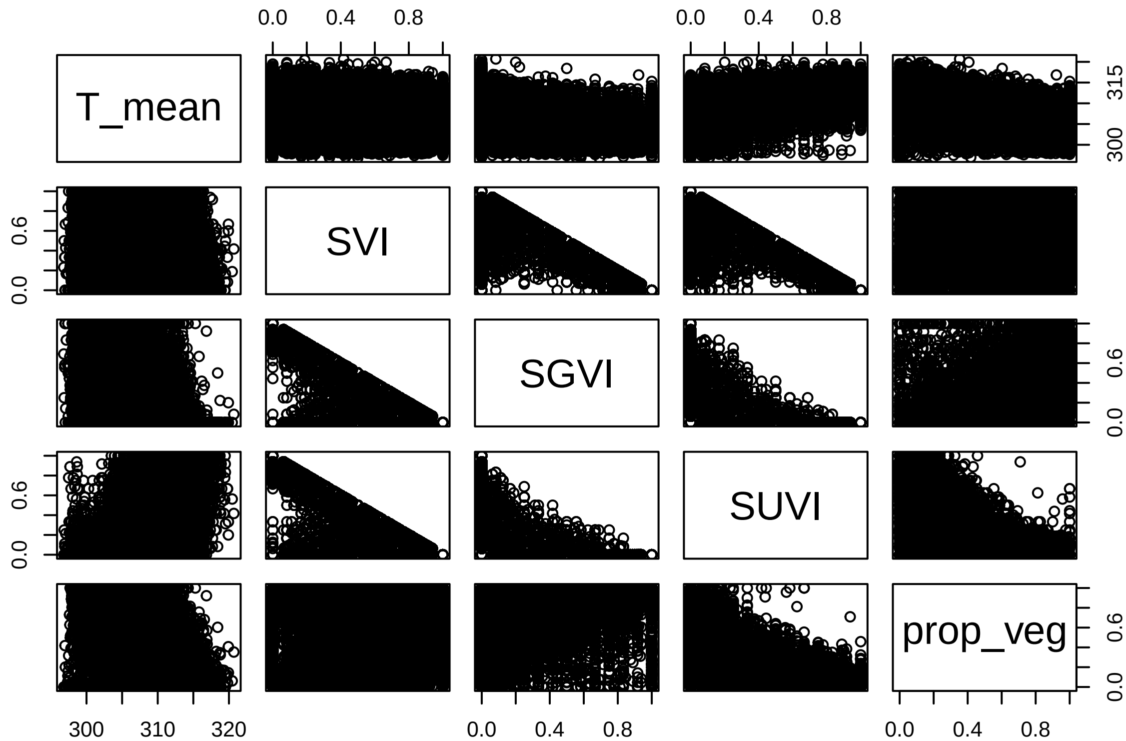
```
##
##   1-1   1-2   2-1   2-2   2-3   3-2   3-3
##  2871 48746   252 54775  6732 30974  3832
```



The variables are highly colinear, as we can see in pairwise correlation

```
predictors <- c("T_mean",
                #"SBmean",
                "SVI",
                "SGVI",
                "SUVI",
                "prop_veg")

pairs(as.data.frame(cdata)[, predictors])
```



whe can see high correlation among predictor variables. Lets check the colective explained variance