

## Nota Técnica: Nova Metodologia do Cálculo do ICMS Verde com aplicação no software R

Letícia de Sousa Borges<sup>1</sup>

Renato Pereira Chaves<sup>2</sup>

Antônio Cordeiro Santana<sup>3</sup>

A nova metodologia de cálculo consiste em simplificar o processo de cálculo do índice ecológico, utilizando técnicas estatística para a composição do mesmo e apresentar a realidade dos municípios paraenses.

A segunda metodologia utiliza a técnica de análise multivariada de dados por análise de fatores, empregando todas as variáveis simultaneamente, além de agrupar essas em fatores por seu grau de correlação. Cada fator tem seu peso que varia todo ano conforme a atualização da base de dados, atentando-se a ordem dos pesos, no qual o fator 1 (Regularização Ambiental) tem peso maior que o fator 2 (Gestão Territorial), este tem peso maior que o fator 3 (Estoque Florestal), possuindo peso maior que fator 4 (Exercício da Gestão Ambiental Municipal), essa distribuição de pesos é garantida pela própria metodologia de cálculo.

A atual metodologia de cálculo, definida pelo Decreto Estadual nº 1.064, de 28 de setembro de 2020, também utiliza a técnica multivariada de dados, porém por componentes principais, permitindo a construção do índice com aplicação simultaneamente de todas as variáveis do modelo assim como a anterior. A principal diferença entre as metodologias é atribuição de pesos nas variáveis simplificando a análise do resultado final.

As variáveis que compõem o modelo são Cadastro Ambiental Rural (CAR), Área de Reserva Legal (ARL), Área de Preservação Permanente (APP), Remanescente de Vegetação Nativa (RVN), Área Antropizada (AA), Área de Uso Restrito (UR), Área de Uso Sustentável (US) e Análise de Cadastro Ambiental Rural Municipal (ACar), devendo essas ser colocadas em proporção das áreas dos municípios, com exceção da última. Ressaltando que UR é a soma das variáveis Terras Indígenas, Áreas Militares e Proteção Integral de esfera Federal, Estadual e Municipal, e US é composto por Territórios Quilombolas, Reserva Pesqueira, Reserva Particular de Patrimônio Natural e Uso Sustentável Federal, Estadual e Municipal.

---

<sup>1</sup> Técnica de Gestão de Meio Ambiente - Estatística – DIORED/SEMAS

<sup>2</sup> Técnico de Gestão de Meio Ambiente – Engenheiro Agrônomo – DIGEO/SEMAS

<sup>3</sup> Professor Titular da Universidade Federal Rural da Amazônia

A variável CAR é oriunda do SICAR/Pará, devendo esta ser colocada na proporção da área cadastrável de cada município, podendo ser representada pela equação 1.

$$pCAR_j = \frac{CAR_j}{CAD_j} \quad j = 1, 2, 3, \dots, 144 \quad (1)$$

em que  $pCAR_j$  é a proporção de área cadastrada para o município  $j$ ,  $CAR_j$  é a área cadastrada do município  $j$  e  $CAD_j$  é a área cadastrável do município  $j$ .

ARL também é oriunda do SICAR/Pará, porém só pode ser medida após a inscrição do cadastro ambiental rural, dessa forma deve ser realizada na proporção do CAR do município assim como APP.

$$pVCAR_{ij} = \frac{vCAR_{ij}}{CAR_j}, \quad \forall i = 1, 2. \quad (2)$$

em que  $pVCAR_{ij}$  é a proporção da variável  $i$  que são obtidas através do SICAR/PA após a inscrição do cadastro ambiental rural para o município  $j$ ,  $vCAR_{ij}$  é a área da variável  $i$  oriunda do SICAR/PA após a inscrição do cadastro ambiental rural para o município  $j$  e  $CAR_j$  é a área cadastrada do município  $j$ .

Outras variáveis que não são oriundas do SICAR/Pará devem ser colocadas em proporção da área do município, sendo as variáveis RVN, AA, UR e US.

$$pVar_{jk} = \frac{Var_{jk}}{AMun_j}, \quad \forall k = 1, 2, 3, 4. \quad (3)$$

em que  $pVar_{jk}$  é a proporção da variável  $k$  que não são obtidas através do SICAR/PA para o município  $j$ ,  $Var_{jk}$  é a área da variável  $k$  oriunda de bases externas ao SICAR/PA para o município  $j$  e  $AMun_j$  é a área da superfície territorial do município  $j$ .

A variável AA após estar em proporção pela área do município deve ser padronizada para computar negativamente no modelo por se tratar de desmatamento, veja a equação 4.

$$pAAp_j = \frac{Max(pAA) - pAA_j}{Max(pAA) - Min(pAA)} \quad (4)$$

em que  $pAAp_j$  é a área antropizada positiva para o município  $j$ ,  $Max(pAA)$  é o valor máximo da área antropizada dos 144 municípios,  $Min(pAA)$  é o valor mínimo da área antropizada dos 144 municípios e  $pAA_j$  é a proporção da área antropizada para o município  $j$ .

Para ACar deve ser computada como uma variável dummy, ou seja, caso o município já esteja habilitado para analisar cadastro ambiental rural terá o valor 1 e os municípios não



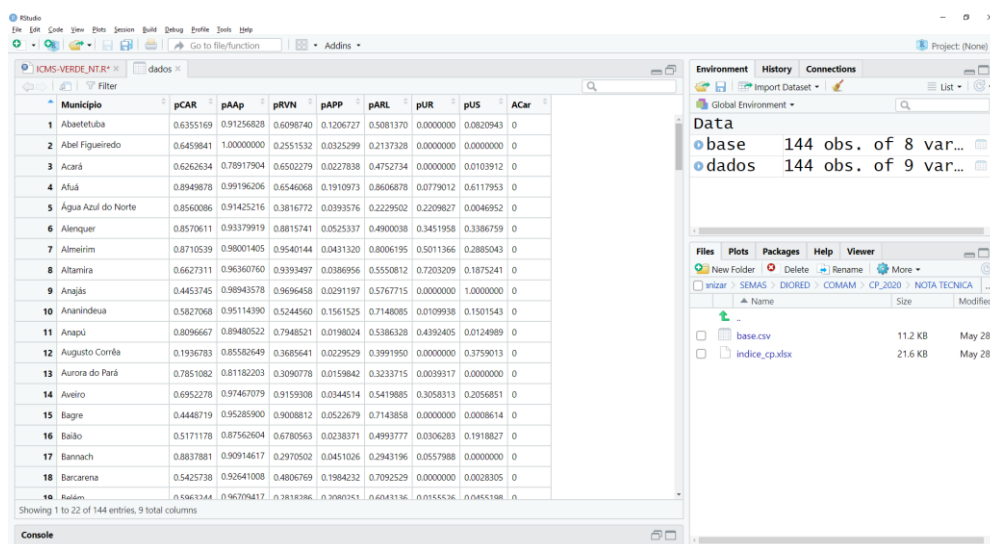
habilitados assumiram valor 0. Dessa forma, os municípios que estão aptos para análise de CAR pontuaram no total do peso dessa variável, porém essa medida deve ser aplicada apenas nos primeiros anos dessa metodologia, pois apenas 14 municípios estão aptos a análise de CAR (Canaã dos Carajás, Novo Progresso, Pacajá, Paragominas, Parauapebas, Redenção, Rondon do Pará, Santana do Araguaia, Santarém, São Félix do Xingu, Tailândia, Tomé-Açu, Ulianópolis e Uruará), servindo de incentivo para os demais municípios. Posteriormente deve ser calculado na proporção das análises realizadas nos municípios, incentivando o aumento das análises municipais. E por fim será feita pela equipe técnica do ICMS Verde uma forma de avaliar a qualidade das análises de CAR's que será aplicada nos próximos anos após as duas outras etapas.

A nova metodologia de cálculo foi realizada pelo software R, por ser gratuito e ter linguagem dinâmica e difundida mundialmente, além de facilitar o acompanhamento do público externo na geração do índice. Para facilitar a criação do algoritmo foi utilizado o pacote *psych*, empregado principalmente para análise multivariada.

A base de dados deve estar salva em arquivo .csv e os dados em proporção como explicado anteriormente. Para exemplificar iremos utilizar a base de dados do índice definitivo do ICMS Verde 2020, vigência em 2021.

Os dados utilizados para exemplificação no software já foram colocados em proporção, como explicado anteriormente e AA já está computando positivamente para o modelo. O código para a leitura é:

```
dados = read.csv2('base.csv', header=TRUE, sep=';', dec='.', stringsAsFactors = FALSE)
base = dados[,-1]
```



Município	pCAR	pAap	pRVN	pAPP	pARL	pUR	pUS	ACar
1 Abaetetuba	0.6355169	0.91256828	0.6098740	0.1206727	0.5081370	0.0000000	0.0820943	0
2 Abel Figueiredo	0.6459841	1.00000000	0.2551532	0.0325299	0.2137328	0.0000000	0.0000000	0
3 Acará	0.6262634	0.78917904	0.6502279	0.0227838	0.4752734	0.0000000	0.0103912	0
4 Açu	0.8949878	0.99196206	0.6546068	0.1910973	0.8606878	0.0779012	0.6117953	0
5 Água Azul do Norte	0.8560086	0.91425216	0.3816772	0.0393576	0.2229502	0.2298227	0.0046952	0
6 Alenquer	0.8570611	0.93379919	0.8815741	0.0525337	0.4900038	0.3451958	0.3386759	0
7 Almeirim	0.8710539	0.96001405	0.9540144	0.0431320	0.8006195	0.5011366	0.2885043	0
8 Altamira	0.6627311	0.96360760	0.9393497	0.0386956	0.5550812	0.7203209	0.1875241	0
9 Anajás	0.4453745	0.98943578	0.9696458	0.0291197	0.5767715	0.0000000	1.0000000	0
10 Ananindeua	0.5827068	0.95114390	0.5244560	0.1561525	0.7148085	0.0109938	0.1501543	0
11 Anapá	0.8096667	0.89480522	0.7948521	0.0198024	0.5386328	0.4392405	0.0124989	0
12 Augusto Corrêa	0.1936783	0.85582649	0.3685641	0.0229529	0.3991950	0.0000000	0.3759013	0
13 Aurora do Pará	0.7851082	0.81182203	0.3090778	0.0159842	0.3233715	0.0039317	0.0000000	0
14 Aveiro	0.6952278	0.97467079	0.9159308	0.0344514	0.5419885	0.3058313	0.2056851	0
15 Bagre	0.4448719	0.95285900	0.9008812	0.0522679	0.7143858	0.0000000	0.0008614	0
16 Baão	0.5171178	0.87562604	0.6780563	0.0238371	0.4993777	0.0306283	0.1918827	0
17 Bannach	0.8837881	0.90914617	0.2970502	0.0451026	0.2943196	0.0557988	0.0000000	0
18 Bercarena	0.5425738	0.92641008	0.4806769	0.1984232	0.7092529	0.0000000	0.0028305	0
19 Belém	0.6043344	0.96709417	0.3818386	0.3080251	0.6043336	0.0155526	0.0855198	0

O segundo passo é verificar a correlação dos dados. Note que “correlacaoP” é significância (P-valor) da correlação entre as variáveis e “Cor” é o grau de correlação.

```

correlacaoP = base
for (i in 1:ncol(base)){
  for (j in 1:ncol(base)) {
    CorP = round(cor.test(base[,j], base[,i], method = "pearson")$p.value,4)
    correlacaoP[j,i] = CorP
  }
}
correlacaoP = round(correlacaoP[1:8,],7)
rownames(correlacaoP) = c("pCAR", "pAAp", "pRVN", "pAPP", "pARL", "pUR", "pUS",
"ACar")
print(correlacaoP)

Cor = cor(base, method = "pearson" )
print(Cor)

```

A terceira etapa é verificar a Esfericidade de Bartlette, Kaiser-Meyer-Olkin (KMO) e Medida de Adequação da Amostra, para realizar esses cálculos devem ser instalado e liberado o pacote psych, veja abaixo.

```

install.packages("psych")
library(psych)
esf = cortest.bartlett (Cor, n = nrow(base))
Sig= esf$p.value
print(paste("Aproximação ao Qui-Quadrado: ",round(esf$chisq,4),
"|Grau de Liberdade: ", esf$df, "| Significância: ", round(esf$p.value,2)))

p.corr <- function(R){
  Rinv <- solve(R)
  Vp <- 1/sqrt(diag(Rinv))
  Vp <- diag(Vp)
  Rp <- -Vp%*%Rinv%*%Vp
}

```

```
diag(Rp) <- 0
Rp
}
calcula.kmo <- function(R){
  Rp <- p.corr(R)
  p <- nrow(R)
  somar <- 0
  somap <- 0
  somar.j <- rep(0,p)
  somap.j <- rep(0,p)
  for(i in 1:p){
    somar <- somar + sum(R[i,]^2)-1
    somap <- somap + sum(Rp[i,]^2)
    somar.j[i] <- sum(R[i,]^2)-1
    somap.j[i] <- sum(Rp[i,]^2)
  }
  kmo <- somar/(somar+somap)
  msa <- somar.j/(somar.j+somap.j)
  cat("\n KMO = ",round(kmo,4),"\n")
  list("kmo"=round(kmo,4),"msa"=round(msa,4),
       "Rp"=round(Rp,3))
}

KMO = calcula.kmo(Cor)
MAA = KMO$msa; MAA
```

O próximo etapa é realização dos cálculos pelas componentes principais obtendo a comunalidade, especificidade e pesos das variáveis. Os parâmetros para o cálculo das componentes principais são: geração de 8 dimensões (o total de variáveis), rotação varimax, cálculo feito pela matriz correlação e o cálculo do escore por regressão.

```
k <- ncol(base)
pc <- principal(r= base, nfactors=k, rotate="varimax", scores=T, covar = FALSE,
               residuals = TRUE, method = "regression")
```

```
cp = pc$rot.mat[,1:k] %*% diag(pc$Vaccounted[1,1:k])
colnames(cp) <- paste("Dimensão", 1:k, sep = " ")
```

Para apresentar a comunalidade, especificidade e variância inicial, utiliza-se a seguinte expressão:

```
comum = round(rowSums(pc$rot.mat^2),1)
esp = round(diag(Cor - comum),2)
est <- cbind(comum, esp, diag(Cor))
rownames(est) <- c("pCAR", "pAAp", "pRVN", "pAPP", "pARL", "pUR", "pUS", "ACar")
colnames(est) <- c("Comunalidade", "Especificidade", "Variância Inicial")
print(est)
```

> print(est)

	Comunalidade	Especificidade	Variância Inicial
pCAR	1	0	1
pAAp	1	0	1
pRVN	1	0	1
pAPP	1	0	1
pARL	1	0	1
pUR	1	0	1
pUS	1	0	1
ACar	1	0	1

Para os pesos das variáveis temos que considerar a seguinte expressão:

$$A_{(8 \times 8)} = \left( \frac{|a_{ml}|}{\sum_{l=1}^8 |a_{l}|} \right); \quad \text{com } m, l = 1, 2, \dots, 8. \quad (5)$$

$$\gamma_{(8 \times 1)} = A_{(8 \times 8)} \times \lambda_{(8 \times 1)}$$

$$\theta_{l(8 \times 1)} = \frac{\gamma_r}{\sum_{r=1}^8 \gamma_r} \times 100; \quad \text{com } r, l = 1, 2, \dots, 8.$$

em que:  $A_{(8 \times 8)}$  é o modulo das cargas (*loadings*) padronizadas,  $a_{ml}$  são os coeficientes da matriz de valores absolutos dos autovetores de cada componentes transformados,  $\sum a_l$  é a soma dos coeficientes relativos da matriz de autovetores das variáveis  $l$ ,  $\gamma_{(8 \times 1)}$  é o vetor



peso não padronizado,  $\lambda_{(8 \times 1)}$  são os autovalores das variáveis  $l$ ,  $\theta_{l(8 \times 1)}$  são os pesos das variáveis.

Note que *pc\$rot.mat* são as cargas (*loadings*), *ajuste* é representado pela a matriz *A* na equação, o pré-peso (os pesos não padronizados em percentual) são representado por  $\gamma$  na equação e *peso* no algoritmo, e “*pesoF*” é o peso final do das variáveis representando  $\theta_i$  na equação 5. Dessa forma temos para a geração dos pesos das variáveis:

```

CAR = round(abs(pc$rot.mat[,1])/sum(abs(pc$rot.mat[1:k,1])),7)
AA = round(abs(pc$rot.mat[,2])/sum(abs(pc$rot.mat[1:k,2])),7)
RVN = round(abs(pc$rot.mat[,3])/sum(abs(pc$rot.mat[1:k,3])),7)
APP = round(abs(pc$rot.mat[,4])/sum(abs(pc$rot.mat[1:k,4])),7)
ARL = round(abs(pc$rot.mat[,5])/sum(abs(pc$rot.mat[1:k,5])),7)
UR = round(abs(pc$rot.mat[,6])/sum(abs(pc$rot.mat[1:k,6])),7)
US = round(abs(pc$rot.mat[,7])/sum(abs(pc$rot.mat[1:k,7])),7)
Acar = round(abs(pc$rot.mat[,8])/sum(abs(pc$rot.mat[1:k,8])),7)
ajuste = matrix(c(CAR, AA, RVN, APP, ARL, UR, US, Acar), ncol = 8, nrow = 8)

peso = ajuste %*% round(pc$Vaccounted[1,c("RC1", "RC2", "RC3", "RC4", "RC5", "RC6",
"RC7", "RC8")],7)
pesoF = round((peso/colSums(peso))*100,7)
pesoF = data.frame(pesoF)
rownames(pesoF)= c("CAR", "AA", "RVN", "APP","ARL", "UR", "US", "ACar")
colnames(pesoF) = c("Pesos")
print(pesoF)

```

```

> print(pesoF)
      Pesos
CAR 13.312543
AA  11.497667
RVN 11.391632
APP 14.473264
ARL 13.478113
UR  12.385268
US  14.450030
Acar 9.011483

```



Para apuração do Índice Geral do ICMS Verde ( $IG_{ICMS\ Verde_j}$ ) foi empregada a soma das multiplicações dos pesos com o valor de cada variável.

$$IG_{ICMS\ Verde_j} = \theta_1 pCAR_j + \theta_2 pAAp_j + \theta_3 pRVN_j + \theta_4 pAPP_j + \theta_5 pARL_j + \theta_6 pUR_j + \theta_7 pUS_j + \theta_8 ACar_j \quad (6)$$

Dessa forma temos,  $l = 1, 2, 3, 4, 5, 6, 7$  e  $8$ , podendo os pesos ser representados por  $\theta_l$ .

em que:  $j$  são 144 municípios,  $l$  são as variáveis que compõe o índice,  $IG_{ICMS\ Verde_j}$  é o índice geral para determinado município  $j$ ,  $\theta_l$  são os pesos das variáveis que compõe o índice,  $pCAR_j$  é a proporção da área cadastrada do município  $j$ ,  $pAAp_j$  é a proporção da área antropizada positiva do município  $j$ ,  $pRVN_j$  é a proporção da área de reserva de vegetação nativa do município  $j$ ,  $pAPP_j$  é a proporção da área preservação permanente do município  $j$ ,  $pARL_j$  é a área de reserva legal do município  $j$ ,  $pUR_j$  é proporção da área de uso restrito do município  $j$ ,  $pUS_j$  é a proporção da área de uso sustentável do município  $j$  e  $ACar_j$  é a análise de CAR realizada pelo município  $j$ .

Em que temos:

$$IG = (\text{base}[, "pCAR"] * \text{pesoF}["CAR",] + \text{base}[, "pAAp"] * \text{pesoF}["AA",] + \text{base}[, "pRVN"] * \text{pesoF}["RVN",] + \text{base}[, "pAPP"] * \text{pesoF}["APP",] + \text{base}[, "pARL"] * \text{pesoF}["ARL",] + \text{base}[, "pUR"] * \text{pesoF}["UR",] + \text{base}[, "pUS"] * \text{pesoF}["US",] + \text{base}[, "ACar"] * \text{pesoF}["ACar",])$$

E por fim para a geração do índice final ( $IF_{ICMS\ Verde_j}$ ) que corresponde ao ajuste do índice geral em 8% (equivalente ao percentual de repasse do ICMS Verde para os 144 municípios).

$$IF_{ICMS\ Verde_j} = \frac{IG_{ICMS\ Verde_j}}{\sum_{i=1}^{144} IG_{ICMS\ Verde}} \times 8 \quad (7)$$

em que  $IF_{ICMS\ Verde_j}$  é o índice final do ICMS Verde para o município  $j$ ,  $IG_{ICMS\ Verde_j}$  é o índice geral do ICMS Verde do município  $j$ , e  $\sum_{i=1}^{144} IG_{ICMS\ Verde}$  é a soma do índice geral dos 144 municípios do Estado.

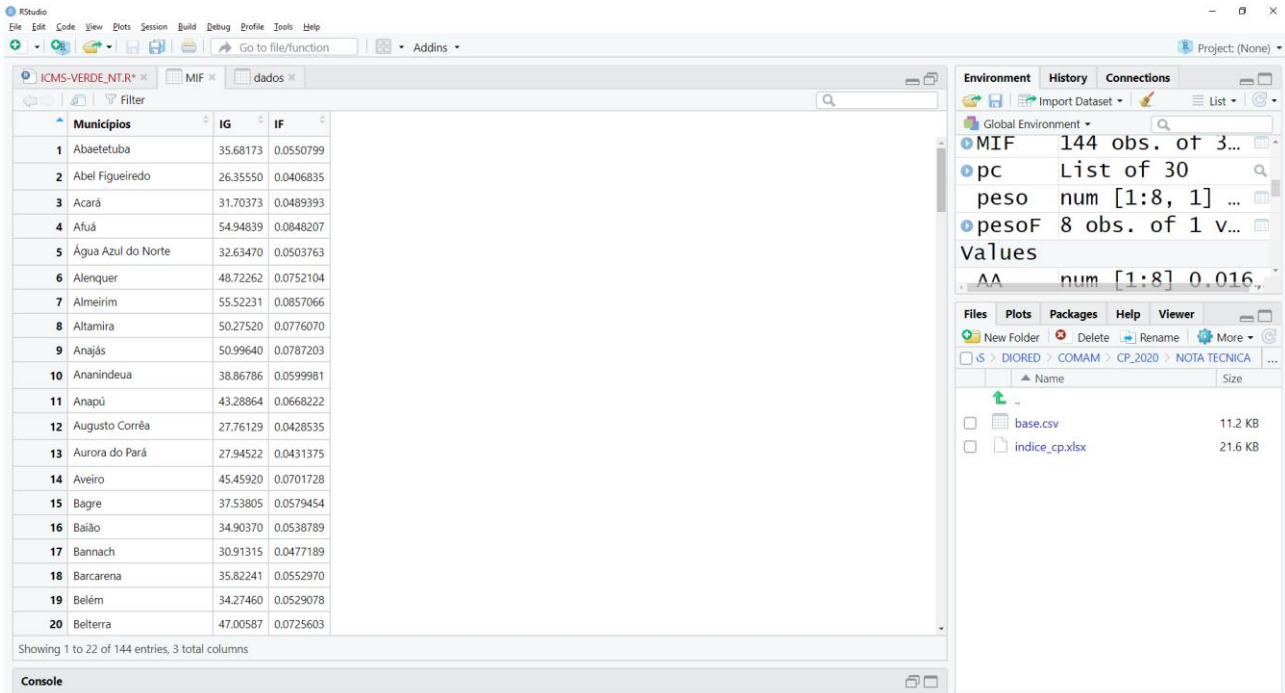
Passando para o software R, temos:

$$IF = \text{round}((IG/\text{sum}(IG)) * 8, 7)$$



E para salvar os resultados do índice geral e final em arquivo .txt, utiliza-se a seguinte expressão:

```
MIF = data.frame(dados[,1],IG, IF)
colnames(MIF) = c("Municípios","IG","IF")
write.table(MIF, file = "Resutados_Pearson_Coeficiente.txt", sep ="\t")
```



The screenshot shows the RStudio interface. The main window displays a data frame with 20 rows and 3 columns: 'Municípios', 'IG', and 'IF'. The environment pane on the right shows the 'Global Environment' with variables 'MIF' (144 obs. of 3...), 'pc' (List of 30), 'peso' (num [1:8, 1] ...), 'pesoF' (8 obs. of 1 v...), and 'Values' (num [1:8] 0.016...).

Municípios	IG	IF
1 Abaetetuba	35.68173	0.0550799
2 Abel Figueiredo	26.35550	0.0406835
3 Acará	31.70373	0.0489393
4 Afuá	54.94839	0.0848207
5 Água Azul do Norte	32.63470	0.0503763
6 Alenquer	48.72262	0.0752104
7 Almeirim	55.52231	0.0857066
8 Altamira	50.27520	0.0776070
9 Anajás	50.99640	0.0787203
10 Ananindeua	38.86786	0.0599981
11 Anapó	43.28864	0.0668222
12 Augusto Corrêa	27.76129	0.0428535
13 Aurora do Pará	27.94522	0.0431375
14 Aveiro	45.45920	0.0701728
15 Bagre	37.53805	0.0579454
16 Baião	34.90370	0.0538789
17 Bannach	30.91315	0.0477189
18 Barcarena	35.82241	0.0552970
19 Belém	34.27460	0.0529078
20 Belterra	47.00587	0.0725603

Belém, 29 de setembro de 2020.