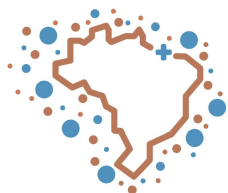


Minicurso:

análise de dados com indicadores reais (Chagas/Leish)



Cenários+

<http://cenarios.unb.br>

/



GOVERNO FEDERAL
UNIÃO E RECONSTRUÇÃO



MINISTÉRIO DA
SAÚDE



UnB



Cenários
e modelagens



SBMT
SOCIETY OF BRASILEAN OF MEDICAL TROPICAL



DOENÇA DE
CHAGAS

CONHECER • PREVENIR • CUIDAR



Saúde nos trópicos:
uma visão integradora
sob a perspectiva do sul global

Medtrop 16a19 de agosto
2026

BRÁSILIA • BRASIL

Minicurso:

análise de dados com indicadores reais (Chagas/Leish)

ENTROPIA MÁXIMA & RISCO EPIDEMIOLÓGICO

O uso da entropia (MaxEnt) na análise de risco da doença de Chagas

Modelagem de nicho ecológico de vetores triatomíneos para mapear áreas de risco de transmissão



O desafio epidemiológico

Por que precisamos de modelos preditivos de risco



A transmissão vetorial depende da distribuição de triatomíneos, difícil de mapear por amostragem direta em toda a área endêmica.



Muitas regiões não são classificadas como “alto risco”, mas registram casos autóctones esporádicos de forma contínua.



Dados de ocorrência são não sistemáticos (coleta oportunística, agentes de vigilância, comunidade) — sem verdadeiros dados de ausência.

Por que a entropia máxima?

Presença-apenas

MaxEnt estima o risco só com dados de presença, sem exigir ausências confirmadas

Sem dados

Prediz para regiões nunca amostradas, a partir de variáveis ambientais

Risco espacial

Converte adequabilidade ambiental em mapas binários de risco de transmissão



O princípio da máxima entropia

A ideia matemática por trás do MaxEnt (Phillips, Anderson & Schapire, 2006)

*“Entre todas as distribuições de probabilidade possíveis sobre a paisagem, escolha a de **máxima entropia** (mais uniforme / com menos suposições) que ainda seja consistente com as condições ambientais observadas nos pontos de ocorrência.”*

1

Sem viés desnecessário

Evita presumir mais do que os dados permitem — a distribuição mais “simples” que respeita as restrições observadas.

2

Restrição = médias ambientais

O modelo exige que os valores médios previstos de cada variável (BIO1, BIO12...) coincidam com a média nos pontos de presença.

3

Pseudo-ausências

Como não há ausências reais, o MaxEnt compara os locais de presença com um plano de fundo (“background”) de pontos aleatórios.



Do dado ao mapa de risco: fluxo de trabalho

Etapas típicas nos estudos de ENM (Ecological Niche Modeling) para vetores de Chagas

1

Ocorrências

Registros georreferenciados de captura do vetor (SINAN, redes de vigilância, coleções)

2

Variáveis ambientais

19+ variáveis bioclimáticas (WorldClim/CHELSA) + altitude, uso do solo, NDVI

3

Calibração MaxEnt

Múltiplas combinações de multiplicador de regularização e classes de resposta (kuenm)

4

Seleção do modelo

Melhor modelo por AICc, taxa de omissão e significância (ROC parcial)

5

Mapa de adequabilidade

Superfície contínua (0–1) convertida em mapa binário de risco (percentil 10)

Métricas de validação-chave: AUC (área sob a curva ROC), taxa de omissão, AICc • **Validação em campo:** Índice Kappa (concordância modelo × observação real)



O software MaxEnt

A ferramenta computacional por trás dos mapas de risco

Ficha técnica

Autoria

Phillips, Dudík & Schapire (AT&T Labs / Princeton / AMNH), desde 2004

Licença

Gratuito e de código aberto para uso não comercial — aplicativo Java standalone

Entrada

Coordenadas de presença do vetor + camadas ambientais em raster (clima, altitude, uso do solo)

Saída

Superfície contínua de adequabilidade (0–1), curvas de resposta, contribuição de variáveis, estatísticas de AUC

Interfaces em R

dismo, ENMeval, kuenm e NicheToolBox — usadas nos 4 estudos revisados nesta apresentação

Por que é o padrão do setor

- Algoritmo mais usado em modelagem de nicho ecológico no mundo, com milhares de citações desde 2006
- Testado e validado em dezenas de espécies de triatomíneos, em 4 continentes
- Interface gráfica simples (point-and-click) e também roda via linha de comando / pacotes em R
- Atualizações contínuas (versão 3.4.1 é a mais usada nos estudos revisados)



Como o MaxEnt funciona por dentro

Os componentes técnicos que o usuário calibra a cada rodada

1

Classes de resposta (feature classes)

Linear, quadrática, produto, hinge (dobradiça) e threshold — moldam a forma da curva de resposta a cada variável

2

Multiplicador de regularização

Controla o quanto o modelo se ajusta aos dados de presença; valores maiores evitam overfitting (0,1 a 10 testados nos estudos)

3

Pontos de background

Amostra aleatória da paisagem usada como “referência”, já que não há dados reais de ausência do vetor

4

Teste de Jackknife

Mede o ganho do modelo com e sem cada variável — identifica quais BIOs realmente importam

O ajuste do modelo, passo a passo:

Presenças + background → combinações de classes e regularização (centenas de modelos candidatos) → seleção pelo menor AICc e menor taxa de omissão → modelo final com curvas de resposta e mapa de adequabilidade



Benefícios do MaxEnt para a análise de Chagas

Por que este software se tornou o padrão nos estudos de vetores triatomíneos

1

Funciona com poucos registros

Gera modelos robustos mesmo com amostras pequenas — essencial para espécies raras ou sub-notificadas (ex.: *T. barberi*, n=21; *R. domesticus*, n=1)

2

Automatiza a busca do melhor modelo

Pacotes como kuenm testam centenas de combinações de variáveis e parâmetros, reduzindo a subjetividade na escolha do modelo final

3

Converte adequabilidade em risco acionável

A superfície contínua (0–1) vira, por um limiar simples (ex.: percentil 10), um mapa binário pronto para orientar a vigilância

4

Aponta as variáveis-chave da transmissão

Curvas de resposta e Jackknife revelam se temperatura, precipitação ou altitude conduzem a distribuição de cada espécie

5

Integra-se à validação de campo

Os mapas de MaxEnt orientam onde amostrar (Hidalgo), fechando o ciclo entre modelo estatístico e confirmação real do vetor

6

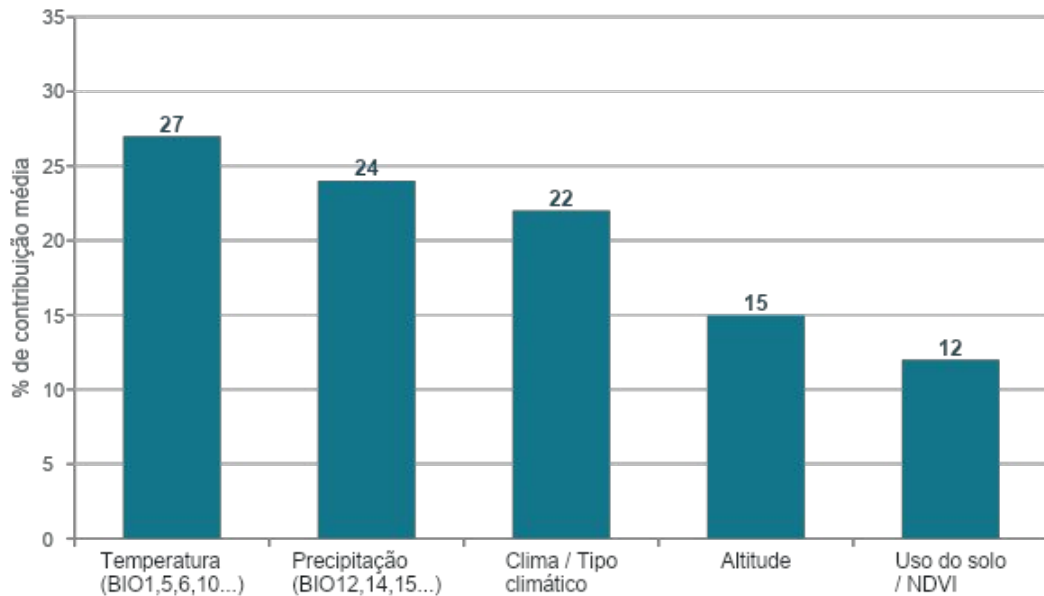
Escalável e comparável entre regiões

O mesmo fluxo foi aplicado no Brasil, Colômbia, México e em toda a América Latina, permitindo comparar risco entre países



Quais variáveis explicam o risco?

Contribuição típica de grupos de variáveis nos modelos revisados



Achados recorrentes

- Sazonalidade da precipitação (BIO15) aparece como fator-chave para várias espécies de *Panstrongylus* e *Triatoma*
- Isotermalidade e temperatura do trimestre mais quente limitam a distribuição em regiões tropicais
- Altitude domina quando há grande variação topográfica (ex.: até 78% de importância em *T. gerstaeckeri*, Hidalgo)
- Clima e uso do solo pesam mais em modelos de escala fina (60 m) que incorporam variáveis antrópicas



Estudo de caso — Espírito Santo, Brasil

Soares et al., 2025 — Spatial and Spatio-temporal Epidemiology

8.507

triatomíneos
coletados (2004–2022)

65

casos de Chagas
registrados (2001–2023)

24

casos autóctones
(13 por vetor)

7

municípios em
área de risco

Modelo de risco: adequabilidade combinada de *P. megistus* (vetor primário) e *T. vitticeps* (principal triatomíneo em ES)

Espécie	AUC	Variáveis mais relevantes
<i>P. díasi</i>	0,969	BIO10 (temp. trim. quente), BIO2
<i>P. tibiamaculatus</i>	0,853	BIO15 (sazon. precipitação), BIO14
<i>P. megistus</i>	0,831	BIO14 (precip. mês mais seco), BIO1
<i>P. geniculatus</i>	0,873	BIO15, BIO4 (sazon. temperatura)
<i>T. vitticeps</i>	0,712	BIO11 (temp. trim. frio), BIO9

Resultado do mapa de risco

A maior parte do estado apresentou alto risco de ocorrência de *P. megistus* / *T. vitticeps*, coincidindo com os pontos de casos autóctones transmitidos por vetor em 7 municípios (ex.: Alfredo Chaves, Santa Teresa, Guarapari).

Primeiro estudo a mapear risco de transmissão vetorial de Chagas no ES.



ELSEVIER

Contents lists available at ScienceDirect

Spatial and Spatio-temporal Epidemiology

journal homepage: www.elsevier.com/locate/sste



Spatial modeling and risk assessment of chagas disease vector distribution in Espírito Santo, Brazil: A comprehensive approach for targeted control

Stefanie Barbosa Potkul Soares^{a,*}, Gustavo Rocha Leite^{b,c},
Guilherme Sanches Corrêa-do-Nascimento^d, Karina Bertazo del Carro^e, Blima Fux^{a,b,*}

^a Infectious Diseases Graduate Program, Federal University of Espírito Santo, Vitória, ES, Brazil

^b Pathology Department, Federal University of Espírito Santo, Vitória, ES, Brazil

^c Animal Biology Graduate Program, Federal University of Espírito Santo, Vitória, ES, Brazil

^d National Institute of the Atlantic Forest, Santa Teresa, ES, Brazil

^e State Department of Health of Espírito Santo, Vitória, ES, Brazil

P. geniculatus	0,075	BIO13, BIO4 (sazon. temperatura)
T. vitticeps	0,712	BIO11 (temp. trim. frio), BIO9

Primeiro estudo a mapear risco de transmissão vetorial de Chagas no ES.

ocorrência de P.
de casos
os (ex.: Alfredo





ELSEVIER

Spatial modeling and r in Espírito Santo, Brazil

Stefanie Barbosa Potkul Soar
Guilherme Sanches Corrêa-do

^a Infectious Diseases Graduate Program, Federal U

^b Pathology Department, Federal University of Esp

^c Animal Biology Graduate Program, Federal Unive

^d National Institute of the Atlantic Forest, Santa Te

^e State Department of Health of Espírito Santo, Vit

P. geniculatus 0,075

T. vitticeps 0,712

Description of the best models selected for each species.

Species	Model Settings	Set of variables	$\Delta AICc$
<i>Panstrongylus diasi</i>	RM: 0.2 FC: qp	Set 4: BIO10, BIO 12, BIO2, BIO3, BIO4	0.0
<i>Panstrongylus tibiamaculatus</i>	RM: 0.1 FC: lqp	Set 36: BIO15, BIO14, BIO4, BIO6, BIO7, BIO1, BIO5	0.0
<i>Panstrongylus megistus</i>	RM: 0.1 FC: q	Set 5: BIO1, BIO5, BIO7, BIO12, BIO13, BIO 14, BIO15	0.0
<i>Panstrongylus geniculatus</i>	RM: 0,9 FC: lqp	Set 28: BIO1, BIO4, BIO5, BIO6, BIO7, BIO12, BIO15	0.0
<i>Triatoma vitticeps</i>	RM: 1 FC: lph	Set 6: BIO1, BIO10, BIO3, BIO5, BIO6, BIO7, BIO8, BIO9, BIO2, BIO11, BIO12, BIO13, BIO14, BIO15, BIO16, BIO17, BIO 18, BIO 19	0.0

RM: regularization multiplier; FC: feature classes (linear = l, quadratic = q, product = p, hinge = h); $\Delta AICc$: delta Akaike information criterion corrected.

ia de P.

fredo

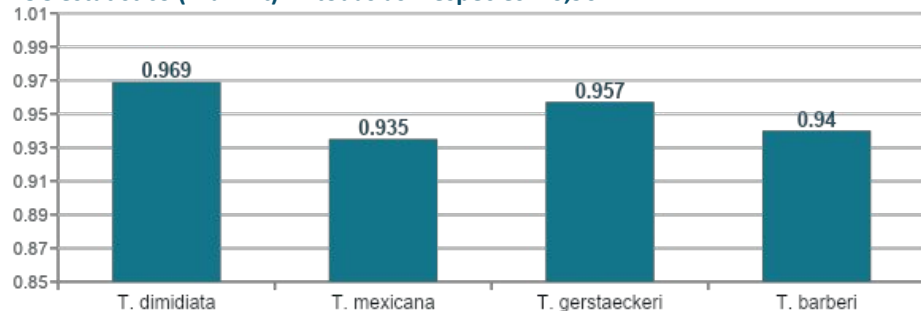
Chagas

Chagas

Estudo de caso— Hidalgo, México

Chico-Avelino et al., 2025, PLOS NTD — primeira validação experimental em campo para Chagas

AUC estatístico (MaxEnt) — todas as 4 espécies > 0,90



Validação de campo — Índice Kappa (concordância modelo × coleta real)

Espécie	Kappa (própria)	Kappa (VCPH)
T. dimidiata	1,00	0,74
T. mexicana	1,00	0,57
T. barberi	0,28	0,19
T. gerstaeckeri	—	0,21

Por que isso importa

- Único estudo revisado que testou o modelo com coleta de campo direcionada, não só dados históricos
- Alto AUC estatístico nem sempre corresponde a alta concordância de campo (ver T. barberi)
- Limiar de corte de 0,75 na probabilidade de presença mostrou-se eficaz para T. dimidiata e T. mexicana
- Reforça a necessidade de validação experimental além das métricas estatísticas do MaxEnt

Estudo

Cl

AUC

1.01

0.99

0.97

0.95

0.93

0.91

0.89

0.87

0.85

0.83

0.81

0.79

0.77

0.75

0.73

0.71

0.69

0.67

0.65

0.63

0.61

0.59

0.57

0.55

0.53

0.51

0.49

0.47

0.45

0.43

0.41

0.39

0.37

0.35

0.33

0.31

0.29

0.27

0.25

0.23

0.21

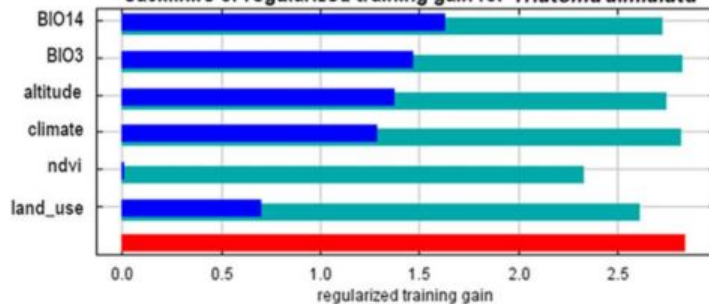
PL

V

Environmental variable

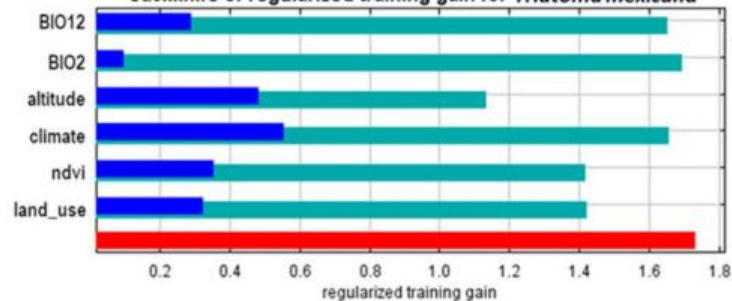
A

Jackknife of regularized training gain for *Triatoma dimidiata*



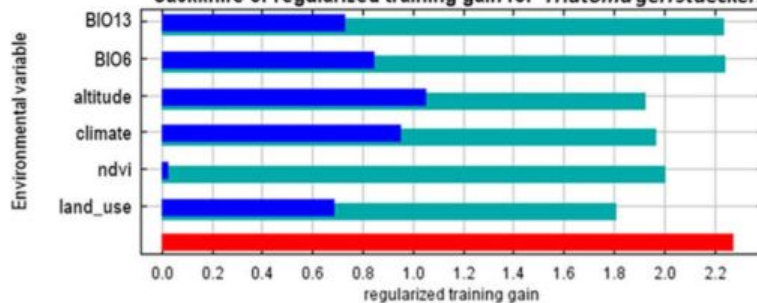
B

Jackknife of regularized training gain for *Triatoma mexicana*



C

Jackknife of regularized training gain for *Triatoma gerrstaeckeri*



D

Jackknife of regularized training gain for *Triatoma barberi*



With only variable



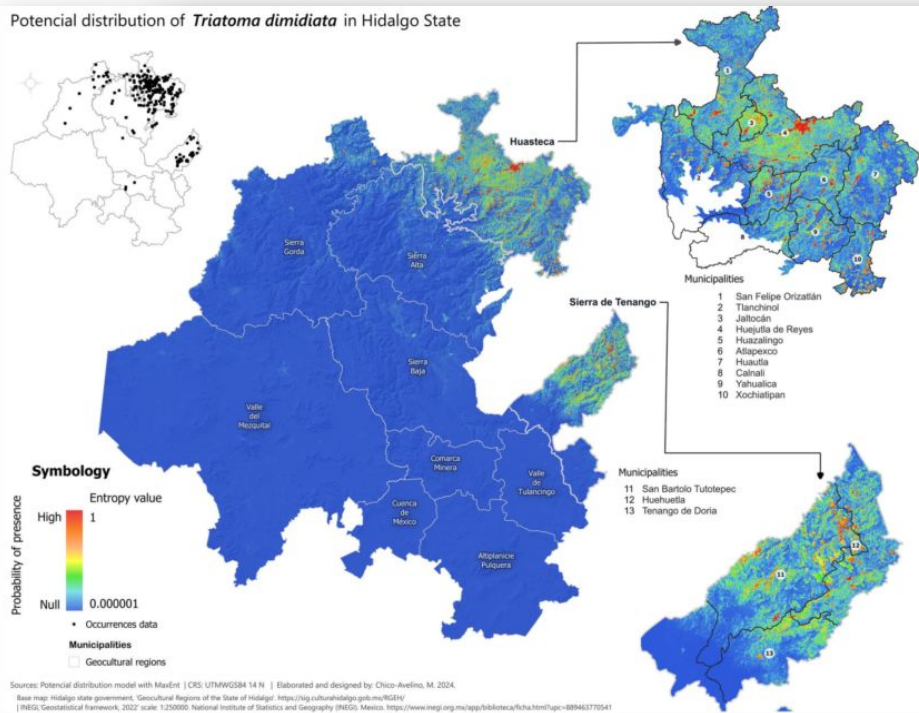
Without variable



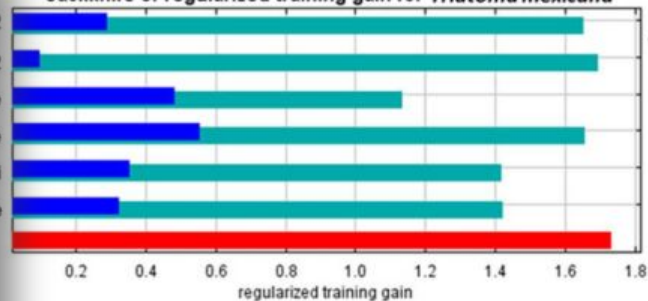
With all variables

* rmanning@cinvestav.mx

Potential distribution of *Triatoma dimidiata* in Hidalgo State



Jackknife of regularized training gain for *Triatoma mexicana*



Jackknife of regularized training gain for *Triatoma barberi*

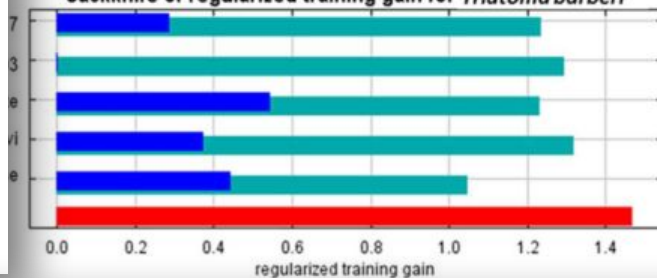


Fig 2. Potential distribution model of *T. dimidiata* in the state of Hidalgo. Colder tones represent the lowest presence probabilities (0.000001) while warmer shades represent the highest values (1). Municipal boundaries are delineated with black borders and identified with white circles containing the municipality name and number listed beside each enlarged section. Lighter grey lines demarcate the geocultural regions. Black dots signify species occurrence records used to train the MaxEnt model. Base map: Hidalgo State Government, Geocultural Regions of the State of Hidalgo (open access). [<https://sig.culturalhidalgo.gob.mx/RGEH/>] | INEGI, 'Geostatistical framework, 2022' scale: 1:250000. National Institute of Statistics and Geography (INEGI), Mexico. [<https://www.inegi.org.mx/app/biblioteca/ficha.html?upc=889463770541>]. This information is also included at the bottom of the map.

<https://doi.org/10.1371/journal.pntd.0013199.g002>

AGRADECIMENTOS



SES
Secretaria de
Estado da
Saúde



Caminhos da inovação para fortalecer a integração da vigilância,
controle e assistência nos Sistemas Universais de Saúde

16 a 19 de agosto

