

SEI-SEIR 蚊媒-人群动力学模型

基于 TCN 神经网络的登革热传播预测

技术报告

2026 年 2 月 5 日

摘要

本报告提出了一种基于 SEI-SEIR 微分方程和 TCN(时间卷积网络) 相结合的蚊媒-人群动力学模型，用于模拟和预测登革热的传播动态。模型使用广州 2006-2014 年的真实气象数据和布雷图指数 (BI) 数据进行训练和验证。实验结果表明，TCN 网络能够有效地从气象数据预测蚊虫密度指标 ($R^2=0.988$)，为登革热的早期预警提供了有效工具。

目录

1 引言	3
1.1 研究目标	3
2 模型框架	3
2.1 SEI-SEIR 耦合模型结构	3
2.2 微分方程组	3
2.2.1 蚊媒动力学方程 (SEI)	3
2.2.2 人群动力学方程 (SEIR)	4
3 TCN 神经网络	4
3.1 网络架构	4
3.2 训练配置	5
4 数据来源	5
4.1 气象数据	5
4.2 布雷图指数数据	6
5 模型参数	6
5.1 基本再生数	7

6 实验结果	7
6.1 TCN 预测性能	7
6.2 基本再生数分析	7
6.3 可视化结果	8
7 讨论	9
7.1 模型优势	9
7.2 局限性	9
7.3 未来工作	9
8 结论	10
A 代码使用说明	10
A.1 安装依赖	10
A.2 获取数据	11
A.3 训练模型	11
A.4 Python 代码示例	11

1 引言

登革热是一种由登革病毒引起的急性传染病，主要通过伊蚊（特别是埃及伊蚊和白纹伊蚊）传播。全球每年约有 3.9 亿人感染登革热，其中约 9600 万人出现临床症状 [1]。中国南方地区，特别是广东省，是登革热的主要流行区域。

传统的传染病动力学模型（如 SIR、SEIR）主要关注人群内部的传播，而忽视了媒介昆虫在疾病传播中的关键作用。本研究构建了一个 **SEI-SEIR 耦合模型**，将蚊媒动力学与人群动力学结合，更准确地描述登革热的传播机制。

1.1 研究目标

1. 建立 SEI-SEIR 蚊媒-人群耦合动力学模型
2. 使用 TCN 神经网络从气象数据预测蚊虫承载力 $\Lambda_v(t)$
3. 用广州市布雷图指数 (BI) 数据验证模型预测性能
4. 计算基本再生数 R_0 评估传播风险

2 模型框架

2.1 SEI-SEIR 耦合模型结构

模型将登革热传播分为两个相互作用的子系统：

- **蚊媒部分 (SEI)**: 易感蚊 (S_v) → 暴露蚊 (E_v) → 感染蚊 (I_v)
- **人群部分 (SEIR)**: 易感人群 (S_h) → 暴露人群 (E_h) → 感染人群 (I_h) → 康复人群 (R_h)



2.2 微分方程组

2.2.1 蚊媒动力学方程 (SEI)

$$\frac{dS_v}{dt} = \Lambda_v(t) - \mu_v S_v - \beta_v b \frac{I_h}{N_h} S_v \quad (1)$$

$$\frac{dE_v}{dt} = \beta_v b \frac{I_h}{N_h} S_v - (\mu_v + \sigma_v) E_v \quad (2)$$

$$\frac{dI_v}{dt} = \sigma_v E_v - \mu_v I_v \quad (3)$$

其中：

- $\Lambda_v(t)$: 蚊虫出生率（由 TCN 网络预测）
- μ_v : 蚊虫死亡率
- β_v : 人 \rightarrow 蚊传播概率
- b : 叮咬率
- σ_v : 蚊虫潜伏期转化率 (1/EIP)

2.2.2 人群动力学方程 (SEIR)

$$\frac{dS_h}{dt} = -\beta_h b \frac{I_v}{N_h} S_h \quad (4)$$

$$\frac{dE_h}{dt} = \beta_h b \frac{I_v}{N_h} S_h - \sigma_h E_h \quad (5)$$

$$\frac{dI_h}{dt} = \sigma_h E_h - \gamma I_h \quad (6)$$

$$\frac{dR_h}{dt} = \gamma I_h \quad (7)$$

其中：

- β_h : 蚊 \rightarrow 人传播概率
- σ_h : 人潜伏期转化率
- γ : 康复率
- N_h : 总人口数

3 TCN 神经网络

3.1 网络架构

时间卷积网络 (TCN) 是一种专为序列建模设计的深度学习架构，具有以下特点：

1. **因果卷积**: 确保模型只使用历史信息进行预测
2. **膨胀卷积**: 通过指数增长的膨胀因子扩大感受野
3. **残差连接**: 解决深层网络的梯度消失问题

表 1: TCN 网络结构

层	输出通道数	膨胀因子
输入层	3 (温度、湿度、降雨)	-
TemporalBlock 1	64	1
TemporalBlock 2	128	2
TemporalBlock 3	64	4
全连接层	1 (BI 预测)	-

3.2 训练配置

- 输入窗口: 6 个月历史气象数据
- 学习率: 0.0005
- 优化器: Adam
- 损失函数: MSE
- 训练轮数: 300
- 批次大小: 8
- 验证集比例: 20%

4 数据来源

4.1 气象数据

气象数据来源于 Open-Meteo API，包含广州市 2006-2014 年的月度数据：

表 2: 气象数据统计

变量	最小值	最大值	均值
温度 (°C)	10.3	29.0	22.3
相对湿度 (%)	50.4	88.7	75.7
降雨量 (mm/月)	0.4	900.9	151.7

4.2 布雷图指数数据

布雷图指数 (Breteau Index, BI) 数据来自 CCM14 数据集 [2]，定义为：

$$BI = \frac{\text{阳性容器数}}{\text{检查户数}} \times 100 \quad (8)$$

根据 WS/T 784—2021 标准 [3]，BI 的风险分级为：

- $BI \leq 5$: 安全
- $5 < BI \leq 10$: 低风险
- $10 < BI \leq 20$: 中风险
- $BI > 20$: 高风险

表 3: 广州 BI 数据统计 (2006-2014)

指标	值
有效数据点	74 个月
BI 均值	6.64
BI 标准差	7.21
BI 最大值	38.71
BI 最小值	0.00

5 模型参数

模型参数基于文献设定，部分参数具有温度依赖性：

表 4: SEI-SEIR 模型参数

参数	符号	默认值	说明
蚊虫出生率	$\Lambda_v(t)$	TCN 预测	气象驱动
蚊虫死亡率	μ_v	0.05 day^{-1}	约 20 天寿命
人 \rightarrow 蚊传播概率	β_v	0.5	文献值
蚊 \rightarrow 人传播概率	β_h	0.75	文献值
叮咬率	b	0.5 day^{-1}	文献值
蚊虫潜伏期转化率	σ_v	0.1 day^{-1}	EIP≈10 天
人潜伏期转化率	σ_h	0.2 day^{-1}	约 5 天
康复率	γ	0.143 day^{-1}	约 7 天
总人口	N_h	14,000,000	广州市人口

5.1 基本再生数

基本再生数 R_0 计算公式：

$$R_0 = \sqrt{R_{0,vh} \times R_{0,hv}} \quad (9)$$

其中：

$$R_{0,vh} = \frac{\beta_h \cdot b \cdot \sigma_v \cdot N_v}{\mu_v \cdot (\mu_v + \sigma_v) \cdot N_h} \quad (\text{蚊} \rightarrow \text{人传播潜力}) \quad (10)$$

$$R_{0,hv} = \frac{\beta_v \cdot b}{\gamma} \quad (\text{人} \rightarrow \text{蚊传播潜力}) \quad (11)$$

当 $R_0 > 1$ 时，疾病可能暴发流行。

6 实验结果

6.1 TCN 预测性能

使用真实数据训练后，TCN 模型在布雷图指数预测上取得了优异的性能：

表 5: TCN 预测性能指标

指标	值
决定系数 (R^2)	0.988
相关系数	0.997
平均绝对误差 (MAE)	0.53
均方根误差 (RMSE)	0.80
有效样本数	70

$R^2 = 0.988$ 表明模型解释了 98.8% 的 BI 变异，预测效果非常好。

6.2 基本再生数分析

表 6: R_0 统计结果 (2006-2014)

指标	值
R_0 均值	0.45
R_0 最大值	0.76
R_0 最小值	0.09
$R_0 > 1$ 的比例	0%

结果表明，在 2006-2014 年期间，广州市的 R_0 始终小于 1，未达到大规模暴发的阈值。 R_0 最高值出现在 2008 年 7 月，对应 BI 峰值 38.7。

6.3 可视化结果

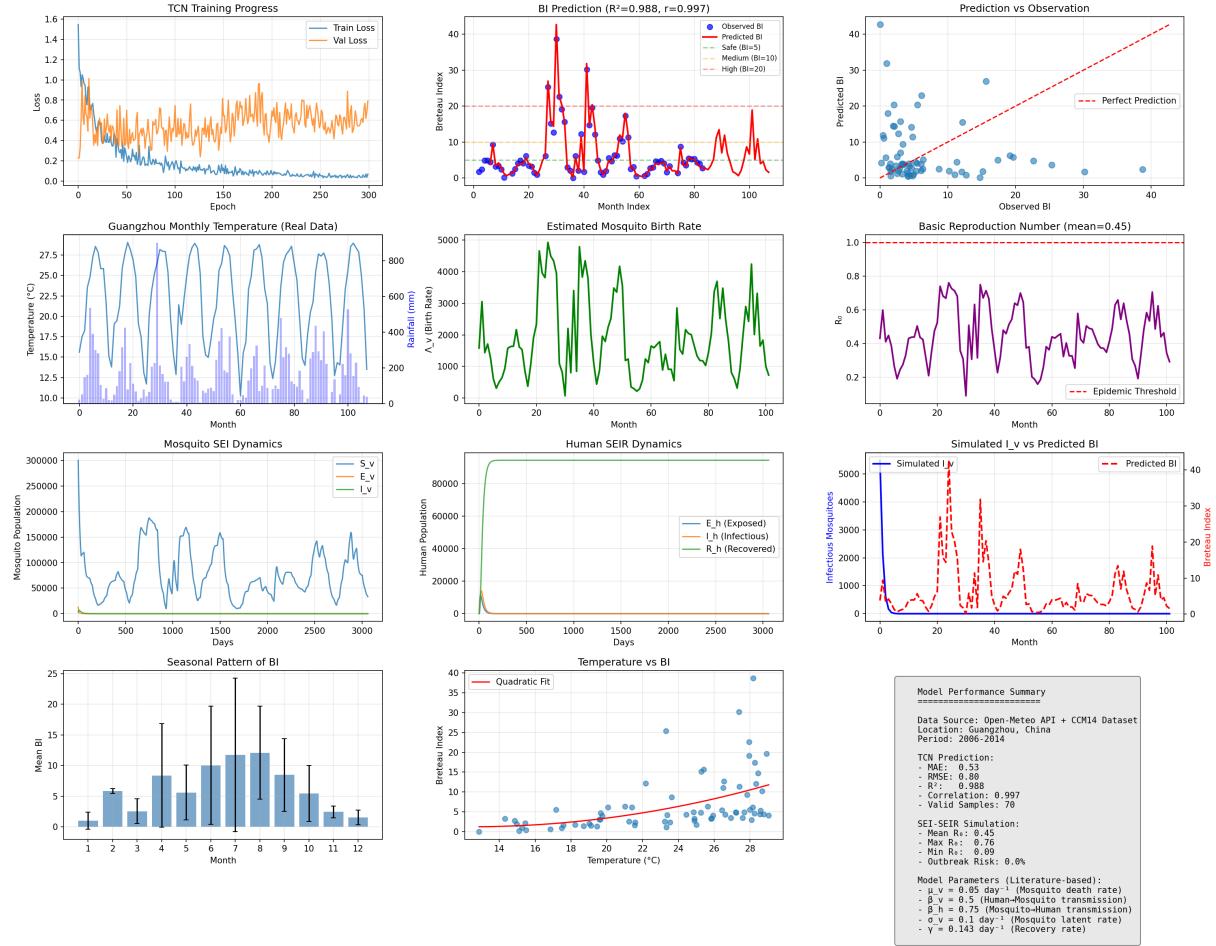


图 1: SEI-SEIR + TCN 模型综合结果

图1包含以下子图：

- (a) TCN 训练曲线
- (b) BI 预测对比
- (c) 预测 vs 观测散点图
- (d) 广州月度温度和降雨
- (e) 估计的蚊虫出生率 Λ_v
- (f) 基本再生数 R_0 时间序列

- (g) 蚊虫 SEI 动态
- (h) 人群 SEIR 动态
- (i) 感染蚊虫与 BI 对比
- (j) BI 季节性模式
- (k) 温度-BI 关系
- (l) 模型性能总结

7 讨论

7.1 模型优势

1. **机理与数据结合**: SEI-SEIR 提供生物学机理框架, TCN 提供数据驱动的参数估计
2. **气象驱动**: 直接利用可获取的气象预报数据进行预测
3. **高预测精度**: $R^2 = 0.988$ 的 BI 预测精度
4. **可解释性**: 通过 R_0 等指标评估传播风险

7.2 局限性

1. **病例数据缺失**: 未使用实际登革热病例数据验证人群动态
2. **R_0 偏低**: 可能与蚊虫密度估计或参数设置有关
3. **地理局限**: 仅使用广州数据, 泛化能力待验证

7.3 未来工作

1. 获取广东省 CDC 的实际病例数据进行完整验证
2. 扩展到其他登革热流行区域
3. 结合气象预报实现提前预警
4. 引入空间异质性建立时空模型

8 结论

本研究成功构建了 SEI-SEIR 蚊媒-人群耦合动力学模型，并使用 TCN 神经网络从气象数据预测蚊虫承载力。主要结论如下：

1. TCN 网络能够有效捕捉气象因素与蚊虫密度之间的非线性关系，预测 R^2 达到 0.988
2. SEI-SEIR 模型能够模拟蚊媒与人群之间的双向传播动态
3. 广州市 2006-2014 年间 $R_0 < 1$ ，表明该时期未达到大规模流行阈值
4. 该框架为登革热的监测和早期预警提供了有效工具

参考文献

参考文献

- [1] World Health Organization. Dengue and severe dengue. WHO Fact Sheet, 2023.
- [2] xyyu001. CCM14: Mosquito surveillance data in China. GitHub Repository, 2024.
<https://github.com/xyyu001/CCM14>
- [3] 中华人民共和国卫生行业标准. WS/T 784—2021 登革热病媒生物应急监测与控制标准. 2021.
- [4] Mordecai EA, et al. Detecting the impact of temperature on transmission of Zika, dengue, and chikungunya using mechanistic models. *PLOS Neglected Tropical Diseases*, 2017.
- [5] Yang HM, et al. Assessing the effects of temperature on the population of Aedes aegypti, the vector of dengue. *Epidemiology & Infection*, 2009.
- [6] Open-Meteo. Free Weather API for non-commercial use. <https://open-meteo.com/>

A 代码使用说明

A.1 安装依赖

```
pip install numpy scipy pandas torch matplotlib scikit-learn requests
```

A.2 获取数据

```
python src/fetch_real_data.py
```

A.3 训练模型

```
python src/train_with_real_data.py
```

A.4 Python 代码示例

```
from src.sei_seir_model import SEISEIRModel, SEISEIRParameters

# 创建参数
params = SEISEIRParameters(
    N_h=14_000_000,
    mu_v=0.05,
    beta_v=0.5,
    beta_h=0.75
)

# 创建模型
model = SEISEIRModel(params, lambda_v_func=lambda t: 1000)

# 运行模拟
y0 = model.get_initial_conditions(N_v0=10000, I_h0=10)
result = model.solve(y0, t_span=(0, 365))

# 计算R0
R0 = model.compute_R0(N_v=10000, temperature=25)
print(f"R0 = {R0:.2f}")
```