arXiv:2602.07170v1 [stat.AP] 6 Feb 2026

用于路线级行程时间可靠性的贝叶斯动态伽马模型

Vadim Sokolov* Refik Soyer†

摘要

路线级行程时间可靠性需要刻画相关路段上总行程时间的分布——现有方法要么假设各路段相互独立（速度快但校准不准），要么通过 Copula 和模拟来建模依赖关系（准确但成本高）。我们提出一种共轭贝叶斯动态伽马模型，引入共同的随机环境来解决这一权衡。给定共享的潜在环境过程时，每个路段的行程时间服从伽马分布；该过程按马尔可夫链演化，在保持条件独立性的同时引入跨路段依赖。通过矩匹配近似，可得路线行程时间的闭式 F 分布，据此可即时计算规划时间指数、缓冲指数和准时概率，且成本与基于独立性的方法相同 $O(1)$。共轭结构确保随着新的传感器数据到达，贝叶斯后验更新和完整预测分布均可用闭式形式得到。将该模型应用于芝加哥 I-55 上覆盖 8.26 英里的 16 个传感器后发现，在计算成本相同的情况下，模型对名义覆盖率为 90% 的预测区间实现了 95.4% 的覆盖率，而基于独立性的卷积方法仅为 34–37%。

1 引言

行程时间可靠性——行程时间的变异性与可预测性——是出行者和交通系统运营者共同关注的首要问题。可靠性不同于平均行程时间：两条平均行程时间相同的路径，其日常变异性可能存在显著差异；研究始终表明，出行者对可靠性的重视程度不亚于，甚至高于对平均行程时间的重视（Liu 等，2004；Tilahun 和 Levinson，2010）。美国联邦公路管理局已将可靠性确定为一项关键绩效指标，并指出，相较于服务水平、速度和流量等其他拥堵指标，可靠性“对公众而言更为重要，也更有意义”（美国联邦公路管理局，2006）。

量化行程时间可靠性，需要超越一阶矩，对行程时间的完整分布进行刻画。这在路径层面尤其具有挑战性，因为

*系统工程与运筹学系，乔治梅森大学。电子邮箱：vsokolov@gmu.edu

†决策科学系，乔治·华盛顿大学。电子邮箱：rsoyer@gwu.edu

总行程时间是多个相关路段行程时间之和，而每个路段行程时间都服从非常规分布。相关的路段到路线聚合问题是一项核心方法学挑战（美国联邦公路管理局，2024）：假设路段之间相互独立的现有方法计算速度快，却会大幅低估路线方差；而对依赖关系进行建模的方法（例如 Copula）则需要耗费大量计算资源进行蒙特卡洛模拟。本文探讨的核心问题是：我们如何才能计算路径级行程时间的预测分布，使其 *(i)* 考虑跨路段依赖，*(ii)* 刻画交通状况动态变化的特性，以及 *(iii)* 得出适用于实时运营的闭式可靠性指标？

我们提出了一种具有共同随机环境的动态伽马分布模型，可同时满足这三项要求。其核心建模思路是，行程时间的跨路段依赖主要由共享条件驱动，例如影响整个走廊的拥堵波、天气和事故。模型通过单一潜在过程η_t来表示这些共享条件，从而在保持条件独立性的同时，在各路段之间形成符合实际的正依赖。这种结构带来了显著的计算优势：无论路段数量多少，路线行程时间分布的计算都可从 m维积分问题简化为单个一维积分。结合伽马矩匹配近似，可得到路线行程时间的闭式 F分布，并可在 $O(1)$时间内计算所有标准可靠性指标——计算成本与简单假设各路段相互独立的方法相同，同时还能正确处理跨路段依赖。

我们的主要贡献如下：

1. 提出一种多元动态伽马模型，通过共同随机环境过程引入跨路段依赖，同时保持解析上的可处理性。该模型将 Nayak（1987）的多元Lomax分布以及 Aktekin 等人（2020）的动态扩展纳为特例。

2. 一种矩匹配近似方法，可将路线行程时间（多个异质伽马随机变量之和）的条件分布简化为单一伽马分布，同时保留共轭结构，并通过 F-分布以闭式形式得到路线预测分布。该方法以与基于独立性的卷积方法相同的 $O(1)$ 计算成本，解决了相关的路段到路线聚合问题。

3. 规划时间指数、缓冲指数和准时到达概率的时变闭式表达式，可随新数据到达而逐步更新。

4. 经验比较表明，由于低估方差，基于独立性的方法对名义覆盖率为 90% 的预测区间仅实现 34–37% 的覆盖率；而我们的模型以相同的计算成本实现了 95.4% 的覆盖率，且校准效果优于基于 Copula 的模拟。

本文其余部分安排如下。第2节回顾行程时间分布、路径级可靠性计算以及交通运输中的动态贝叶斯模型相关文献。

第3节介绍单一路段的动态伽马模型。第4节提出包含共同随机环境的多变量扩展。第 5 节通过矩匹配推导主要理论结果：闭式路径级可靠性。第6节讨论贝叶斯推断。第7节展示对 I-55 交通数据的经验应用。第7节总结全文，并讨论后续扩展。

2 文献综述

2.1 行程时间分布建模

可靠性分析的前提是为行程时间建立恰当的分布模型。早期研究假设行程时间服从正态分布（Iida，1999），这种假设虽能简化路径层面的卷积计算，却无法刻画经验行程时间数据所具有的正支撑、右偏和重尾特征。

经验研究表明，偏态的正值分布更适合描述行程时间。Susilawati et al.（2013）比较了多种用于城市道路行程时间的参数族——伽马分布、对数正态分布、Burr 分布和对数逻辑分布，发现对于较长的链路，对数正态分布和 Burr 分布拟合效果最佳，而伽马分布对于较短的路段也具有竞争力。Fosgerau 和 Fukuda（2012）采用对数正态分布对城市道路上的行程时间进行建模，并将其参数与行程时间变异性价值联系起来。Zang et al.（2022b）的综合综述汇总了已应用的各种分布，并指出没有任何单一参数族能在所有情境下占据优势。

该领域文献中的一个重要观察是区分静态 和动态分布模型。现有大多数研究都将单一分布拟合于合并的行程时间样本，隐式地假设平稳性。然而，行程时间分布会在一天之内发生变化（例如，早晨处于自由流状态，而下午出现拥堵），也会跨天变化（例如，受到事故或天气的影响）。van Lint等人（2008）指出，基于方差的不可靠性度量是不充分的，因为相同的方差可能源自截然不同的分布形状，并主张刻画完整的时变分布。Zang et al.（2022a）近期开展的综合性综述进一步印证了这一观点，将相关文献归纳为一个四阶段框架——特征刻画、评估、价值评估和分配——并指出需要能够同时处理非平稳性与跨路段依赖的模型。

另一个关键区别在于，模型是针对路段层级运行，还是直接对路径级行程时间进行建模。路段级模型更为基础：每个道路路段都有独特的物理特征（长度、通行能力、汇入／汇出几何形状），这些特征会产生各自特有的分布形状。能够捕捉这些路段级动态的模型可以组合起来，对任意路径进行路径级预测；而直接根据路径数据拟合的模型则局限于单一起讫点对。我们的方法通过具有路段专属速率参数 λ_j 的伽马分布，分别对每个路段建模，并基于这些部件构建路径级可靠性。

2.2 路径级可靠性与卷积问题

计算路径行程时间 $S_t = \sum_{j=1}^{m} y_{jt}$ 的分布，需要对路段级分布进行卷积。假设各路段行程时间服从正态分布且彼此独立，则卷积结果显然仍为正态分布（Iida，1999）。对于更符合实际的分布，卷积则没有封闭形式。

Ng 和 Waller（2010年）通过使用快速傅里叶变换（FFT）对卷积积分进行数值求解来应对这一问题，该方法适用于任何具有已知特征函数的分布。Ng 等人（2011）提出了一种基于概率不等式的无分布方法，完全避免指定参数族，但得到的是保守界而非精确分布。Nie 等人（2012）将基于快速傅里叶变换的卷积应用于芝加哥高速公路网络的实用路径引导系统。Lei等人（2014）通过将GEV分布和广义帕累托分布拟合到路段级单位距离行程时间，再对这些分布进行卷积，从而对城市快速路上的行程时间进行建模，以获得不同服务水平下的路径级可靠性。Ramezani和 Geroliminis（2012）采用另一种方法，使用马尔可夫链对每条链路的状态转移进行建模，并通过各链路特定分布的卷积来构建路径行程时间分布。

一份近期的FHWA方法论报告（美国联邦公路管理局，2024）讨论了沿用户定义路径估计旅行时间分布的问题，并指出：在保持相关性的同时由链路级数据构建路径分布，仍是一个核心方法学挑战，尤其是在各机构从固定点传感器数据转向基于轨迹的探测数据之际。

所有这些基于卷积的方法都假设路段行程时间相互独立——这一假设被普遍认为不切实际。路径上相邻路段的交通状况相似，经验数据也表明，连续路段行程时间之间存在很强的正相关性（Zockaie 等人，2014年）。忽略这种依赖关系会低估路线行程时间的方差，并导致预测区间不可靠。Filipovska 和 Mahmassani（2021）针对这一问题，在具有广义时空相关性的随机时变网络中估计路径行程时间分布，利用蒙特卡洛模拟对相关链路行程时间进行卷积，并在行程途中自适应地更新路径分布。近期，Kamphuis 等人（2025）为具有空间依赖性的网络中的逐边行程时间联合分布开发了一种非参数估计量，并证明了其一致性保证；然而，他们的高斯框架无法捕捉行程时间数据固有的偏度。

基于Copula的模型为独立于边缘分布指定路段间依赖提供了一个自然的框架。Chen等人（2017）提出了一种基于联结函数的方法，用于估计城市干道上的行程时间可靠性：他们为各个路段的行程时间拟合边缘分布，然后使用Copula对路段间依赖结构进行建模。他们的方法正确地指出，每个路段都需要各自的分布模型，而且路段间依赖关系十分重要；然而，在其Copula模型下计算路线总和 S_t的分布，需要进行 m 维数值积分或蒙特卡洛模拟，随着路段数量增加，这种方法的扩展性较差。

在网络层面，Saedi等人（2020）使用网络划分，将大型异质网络分解为同质子网络，并利用线性均值-标准差关系估计各分区内的行程时间可靠性。他们的方法为处理网络异质性提供了一种实用方案，但依赖基于仿真的估计，且无法提供闭式路径级分布。

Westgate等人（2016）开发了一种用于救护车路径规划的大规模网络行程时间分布模型，该模型利用对数高斯Cox过程捕捉时空依赖性，但需要采用计算成本高昂的基于仿真的推断。基于仿真的方法适用于离线分析，但不适合实时可靠性计算，因为每当新数据点到达时，都必须更新相关指标。

我们采用不同的机制解决卷积问题：我们不通过Copula对依赖关系建模后再对联合分布积分，而是利用共同随机环境隐式地诱导依赖关系。由此形成的条件独立结构意味着，无论路段数量多少，路线分布都只需对该环境进行一维积分。与Lei等人（2014）和Ramezani和Geroliminis（2012）的卷积方法不同，我们的方法能够自然地处理路段间的依赖关系；与Chen等人（2017）的Copula方法不同，我们的方法避免了 m维积分。

2.3 交通动态与贝叶斯模型

允许参数随时间演变的动态模型适用于交通数据，因为交通状况变化迅速。West和Harrison（1997）的贝叶斯动态线性模型框架提供了一种通用方法，Tebaldi等人（2002）则将其应用于来自北卡罗来纳州研究三角区的高速公路交通流数据。

一种用于刻画时变可靠性的主要方法是源自金融计量经济学的GARCH（广义自回归条件异方差）模型族。Yang和Wu（2018）构建了一个ARIMA-GARCH模型，可为干道行程时间生成时变置信区间：拥堵时区间变宽，自由流状态下则变窄。GARCH框架明确对波动聚集进行建模——即高方差时期往往紧随其他高方差时期出现的现象——这是交通动态的自然特征。然而，GARCH模型针对单一的汇总时间序列（通常是路线级）进行建模，既不提供路段级分布模型，也没有将各部件组合成路径分布的机制。相比之下，我们的折扣因子 γ 发挥着与GARCH持久性参数类似的作用，但它是在共轭贝叶斯框架内运作，能够得到闭式后验更新和完整的预测分布，而不仅仅是置信区间。

Mihaylova 等人（2007）将粒子滤波方法应用于交通状态估计，在宏观交通模型中为高速公路交通流开发了一种粒子滤波器。Polson和Sokolov（2015）开发了用于交通流的贝叶斯粒子跟踪方法，将物理交通模型与统计估计相结合。Carvalho等人（2010）的粒子学习框架

提供了一种基于充分统计量进行序贯贝叶斯更新的通用方法，可直接应用于我们的模型。Kim and Ye（2022）针对稀疏探测数据下的高速公路行程时间估计，开发了一种贝叶斯混合模型，并利用历史相邻数据构建数据驱动先验；他们采用的吉布斯采样方法可得出可信区间，但需要进行MCMC计算，且无法扩展至路径级聚合。

近年来，深度学习方法在交通预测任务中取得了出色的表现。Polson and Sokolov（2017年）展示了深度神经网络能够捕捉短期交通流预测中的复杂非线性模式。然而，深度学习方法通常只能生成点预测，或需要额外校准才能提供可靠的不确定性估计。它们无法直接生成可靠性分析所需的分布建模输出（累积分布函数、分位数、预测区间）。

我们的模型在这一领域中占据独特的位置：对于固定超参数（α，γ，λ），它是一个共轭贝叶斯模型，能够对环境 η_t 进行精确的顺序更新，并以闭式形式给出完整的预测分布。超参数本身未知时，可采用粒子学习方法进行联合在线估计（第 6 节）。通过矩匹配近似，即使在路径层面也能保留共轭性。正如 Zang et al.（2022a）的近期综述所指出的，分布建模与实时计算之间的差距仍是该领域面临的一项关键挑战；我们的闭式方法直接应对了这一挑战。

2.4 共同随机环境模型

在可靠性理论中，利用共享潜变量在条件独立的部件之间引入依赖关系，这一思想由来已久。Lindley and Singpurwalla（1986）提出了由多个部件共享的“共同环境”概念，并表明服从伽马分布的随机环境会使部件寿命呈现多元Lomax分布。Nayak（1987）进一步研究了由此得到的多元Lomax分布的统计性质，包括其矩和依赖结构。Singpurwalla（1995年）将该框架拓展至运行条件随时间变化的动态环境；Hougaard（2000）则全面论述了生存分析中相关的脆弱性模型。

Aktekin 等人（2013）将 Lindley–Singpurwalla 框架拓展至动态贝叶斯框架，其中潜在环境遵循贝塔驱动马尔可夫过程，并将该模型用于抵押贷款违约计数。Aktekin 等人（2020）提出了一类通用的多元非高斯时间序列模型，其中包含伽马、泊松和二项情形；该模型族中的伽马成员可导出本文采用的动态多元贝塔素分布。Aktekin 等人（2020）还提供了第 6 节所用的 FFBS 后向抽样和MCMC推断框架。Soyer et al.（2016）利用共同随机环境结构对计数数据时间序列中的依赖关系进行建模。

我们的论文沿着这一研究方向，进一步拓展到一个具体且具有重要实际意义的领域：我们表明，共同环境结构与针对

路线级求和的矩匹配近似相结合，可得出闭式路径行程时间分布。这一结果将公共随机环境模型从路段级建模框架转变为用于路线级可靠性分析的完整工具。

2.5 研究缺口与贡献

表1总结了现有方法如何满足路径级可靠性分析的关键要求。基于卷积的方法（Ng 和 Waller，2010年；Lei 等人，2014年；Ramezani 和 Geroliminis，2012年）利用路段模型构建路径分布，但假设路段之间相互独立，并采用静态分布。Chen等人（2017）的Copula方法能够正确建模路段级分布及跨路段依赖，但需要进行 m维模拟才能计算路径可靠性，且无法实现动态更新。网络级方法（Saedi等人，2020）完全绕过路段级建模，因此无法为任意路线组合可靠性估计。此前的方法均无法同时提供：（i）路段级分布模型；（ii）能够追踪时变条件的动态更新；（iii）能够捕捉跨路段依赖的多变量结构；以及（iv）闭式路径级可靠性指标。我们的框架弥补了这一缺口。

表1：所提出框架相对于现有路径级行程时间可靠性方法的定位。

方法	动态	依赖关系	闭式
正态卷积（Iida，1999）	No	No	Yes
快速傅里叶变换卷积（Ng 和 Waller，2010年）	No	No	数值方法
GEV/GPD 卷积（Lei 等人，2014）	No	No	数值方法
马尔可夫链（Ramezani 和 Geroliminis，2012）	No	No	数值方法
Copula（Chen 等人，2017）	No	Yes	模拟
相关蒙特卡洛模拟（Filipovska 和 Mahmassani，2021）	Yes	Yes	模拟
网络分区（Saedi等人，2020）	No	隐式	模拟
ARIMA-GARCH（Yang and Wu，2018年）	Yes	No	区间
深度学习（Polson and Sokolov，2017年）	Yes	隐式	No
模拟（Westgate 等，2016）	Yes	Yes	模拟
本文	**Yes**	**Yes**	**Yes**

3 动态 Gamma 模型

3.1 观测模型

给定道路路段的行程时间存在多种不确定性来源，包括每日需求变化或特殊事件导致的需求变化（Clark 和 Watling，2005）、事故和道路施工导致的车道封闭（Chen 等人，2003），以及天气造成的性能下降（Tu 等人，2007）。我们将这些因素分解为可观测协变量和潜在环境分量。

我们假设，道路路段在时段 t 的行程时间 y_t 服从形状参数为 α 、率参数为 $\eta_t e^{\beta^T u_t}$的伽马分布，其中 u_t 是表示环境变量的可观测协变量向量

而 η_t 是时间 t的潜在环境：

$$y_t \mid \alpha, \eta_t, \beta, u_t \sim \mathrm{Gam}(\alpha,\, \eta_t e^{\beta^T u_t}). \tag{1}$$

这里采用率参数化，因此概率密度为

$$p(y_t \mid \alpha, \eta_t, \beta, u_t) = \frac{(\eta_t e^{\beta^T u_t})^\alpha}{\Gamma(\alpha)}\, y_t^{\alpha-1}\, e^{-\eta_t e^{\beta^T u_t} y_t}.$$

平均行程时间为 $E[y_t \mid \eta_t] = \alpha/(\eta_t e^{\beta^T u_t})$，与环境质量 η_t 成反比： η_t 的值越高，条件越好（行程时间越短）；值越低则表示拥堵。

伽马分布是行程时间建模的自然选择。其定义域为 $(0,\ \infty)$，能够适应右偏特征，并且包含指数分布 $(\alpha = 1)$ 这一特例。实证研究证实，伽马分布能够很好地拟合路段级行程时间数据（ Susilawati et al., 2013; Zang et al., 2022b）。

在给定 α 和 $\eta_t e^{\beta^T u_t}$ 的条件下，假设不同时段的行程时间相互独立。我们定义 $D^{t-1} = (y^{t-1},\ u^{t-1})$，其中 $y^{t-1} = (y_1,\ \ldots,\ y_{t-1})$ 和 $u^{t-1} = (u_1,\ \ldots,\ u_{t-1})$。

3.2 动态随机环境

潜在环境 η_t 按照马尔可夫过程演化。给定历史记录 D^{t-1}，我们规定

$$\eta_t = \frac{\eta_{t-1}}{\gamma}\, \epsilon_t, \tag{2}$$

其中 $\epsilon_t \mid \alpha, \gamma, \beta, D^{t-1} \sim \mathrm{Beta}[\gamma a_{t-1}, (1-\gamma) a_{t-1}]$ 且 $0 < \gamma < 1$ 为折扣因子。这种演化方式源自 Singpurwalla（1995年）的动态环境模型，并已用于可靠性增长试验（Lindley and Singpurwalla，1986）和抵押贷款违约率建模（Aktekin 等，2013）。

折扣因子 γ 控制模型适应新信息的速度。 γ 取值较小（快速折扣）时，模型对近期观测值反应灵敏；取值接近 1 时，则会更加重视历史数据。这种演化意味着随机序关系 $\eta_t < \eta_{t-1}/\gamma$ ，并具有随机游走性质：

$$E(\eta_t \mid \eta_{t-1}, D^{t-1}) = \eta_{t-1},$$

其中 $(\eta_t \mid \eta_{t-1}, D^{t-1})$ 是定义在 $(0, \eta_{t-1}/\gamma)$ 上的缩放贝塔密度。

3.3 序贯贝叶斯更新

该模型的一个关键特征是， η_t 的后验分布仍属于伽马分布族，从而能够进行精确的顺序更新。

命题 1。 如果 $(\eta_0 \mid D^0) \sim$ 伽马分布(a_0, b_0)，那么对所有 $t \geq 1$:

(i) 时刻 $t-1$ 的后验分布为 $(\eta_{t-1} \mid D^{t-1}) \sim$ 伽马分布(a_{t-1}, b_{t-1})。

(ii) 时刻 t 的先验分布为 $(\eta_t \mid D^{t-1}) \sim$ 伽马分布$(\gamma a_{t-1}, \gamma b_{t-1})$。

(iii) 观测到 y_t后，后验分布更新为 $(\eta_t \mid D^t) \sim$ 伽马分布(a_t, b_t) 其中

$$a_t = \gamma a_{t-1} + \alpha, \qquad b_t = \gamma b_{t-1} + y_t e^{\beta^T u_t}. \tag{3}$$

证明。 第(ii)部分由马尔可夫演化(2)（Aktekin 等，2013）得出。对于(iii)，伽马似然与(ii)中的Gamma先验共轭；将核相乘得到形状参数 $\gamma a_{t-1} + \alpha$和率参数 $\gamma b_{t-1} + y_t e^{\beta^T u_t}$。

□

更新是直接的： η_t 的共轭Gamma先验与伽马似然相结合，得到伽马后验分布。折扣因子 γ 会系统性地使先验分布相对于后验分布的方差膨胀——陈述(ii)表明，先验方差 $\gamma a_{t-1}/(\gamma b_{t-1})^2 = a_{t-1}/(\gamma b_{t-1}^2)$是后验方差 a_{t-1}/b_{t-1}^2 的 $1/\gamma$倍。

推论 1 (随机性质)。 环境过程满足：

$$E(\eta_t \mid D^{t-1}) = E(\eta_{t-1} \mid D^{t-1}), \qquad V(\eta_t \mid D^{t-1}) > V(\eta_{t-1} \mid D^{t-1}).$$

也就是说，环境的预测均值保持不变，但两次更新之间的不确定性会增加。

3.4 预测分布

通过对环境进行积分消去，可得到 y_t 的一步预测分布：

$$p(y_t \mid \alpha, \gamma, \beta, D^{t-1}) = \int_0^\infty p(y_t \mid \alpha, \eta_t, \beta, u_t)\, p(\eta_t \mid D^{t-1})\, d\eta_t. \tag{4}$$

由此得到复合伽马分布：

$$p(y_t \mid D^{t-1}) = \frac{\Gamma(\alpha + \gamma a_{t-1})}{\Gamma(\alpha)\,\Gamma(\gamma a_{t-1})} \frac{(e^{\beta^T u_t})^\alpha\, y_t^{\alpha-1}\, (\gamma b_{t-1})^{\gamma a_{t-1}}}{(\gamma b_{t-1} + y_t e^{\beta^T u_t})^{\alpha + \gamma a_{t-1}}}, \tag{5}$$

这是一个 Beta prime（复合伽马分布）分布；当 $\alpha = 1$ 时，它简化为 Lomax（Pareto II 型）分布。一个重要结果是，预测分位数可通过 F 分布以闭式形式求得˜：若记 $\tilde{a} = \gamma a_{t-1}$ 和 $b = \gamma b_{t-1}$, 则

$$\frac{\tilde{a}}{\alpha} \cdot \frac{y_t e^{\beta^T u_t}}{\tilde{b}} \;\sim\; F(2\alpha,\, 2\tilde{a}). \tag{6}$$

这意味着，预测累积分布函数、分位数和矩均可使用标准F-分布函数计算，无需进行数值积分或模拟。

注 1. 为使符号简洁，暂不考虑协变量（即设定 $\beta = 0$），预测均值和方差为

$$E[y_t \mid D^{t-1}] = \frac{\alpha\,\tilde{b}}{\tilde{a}-1} \quad (\tilde{a}>1), \qquad V[y_t \mid D^{t-1}] = \frac{\alpha\,\tilde{b}^2(\alpha+\tilde{a}-1)}{(\tilde{a}-1)^2(\tilde{a}-2)} \quad (\tilde{a}>2).$$

存在协变量时，两个表达式分别由 $e^{-\beta^T u_t}$ 和 $e^{-2\beta^T u_t}$缩放。与伽马分布相比，厚尾现象（预测方差被放大$(\alpha+\tilde{a}-1)/(\tilde{a}-2)$倍）反映了关于环境的额外不确定性。

4 多元扩展：共同随机环境

4.1 模型设定

考虑一条由 m 个路段组成的路径。时刻 t路段 j 上的行程时间服从

$$y_{jt} \mid \lambda_j, \alpha, \eta_t, \beta, u_t \sim \mathrm{Gam}(\alpha,\, \lambda_j \eta_t e^{\beta^T u_t}), \tag{7}$$

其中， $\lambda_j > 0$ 是用于描述路段 j相对速度的路段特定参数，而 η_t 是所有路段共享的共同随机环境。 λ_j 较高的路段，其期望行程时间较短： $E[y_{jt} \mid \eta_t] = \alpha/(\lambda_j \eta_t e^{\beta^T u_t})$。

关键的建模假设是**条件独立性**：给定 η_t 和各个 λ_j，路段行程时间 $y_{1t}, \ldots, y_{mt}$ 彼此独立。共同环境 η_t 引入了所有跨路段的依赖关系。这一假设有物理依据：相邻路段之间相关性的主要来源是共享路网状况（拥堵波、天气、事故），而这正是 η_t 所刻画的。

这种依赖结构通过一个共同潜变量使条件独立的观测值产生相关性，其起源可追溯至 Lindley and Singpurwalla（1986）和 Nayak（1987）的共享环境模型，后来被应用于信用风险建模（Aktekin 等，2013年）和计数数据时间序列（Soyer et al., 2016）。

一个重要的结构推论是，任意两个路段之间的预测相关 $i \neq j$ 具有 $\mathrm{Corr}(y_{it}, y_{jt} \mid D^{t-1}) = \alpha/(\alpha+\tilde{a}-1)$ 的形式，它依赖于 α 和 $\tilde{a}$，而不依赖于 i 或 j。也就是说，该模型意味着**等相关**：无论空间间隔如何，所有路段对都具有相同的相关。经验上，相邻路段往往比相距较远的路段表现出更高的相关（第7节）。尽管如此，该模型捕捉了主导的一阶效应——由共享条件驱动的普遍正向相依——而路径级聚合（对所有路段对求和）由平均相关而非其空间结构所支配。我们将在第8节重新讨论这一局限。

4.2 顺序更新

单变量模型的所有结果在修改递推关系后仍然成立。在给定全部 m 个路段的条件下， η_t 的后验分布更新为

$$(\eta_t \mid \lambda, \alpha, \gamma, \beta, D^t) \sim \mathrm{Gam}(a_t, b_t), \tag{8}$$

其中

$$a_t = \gamma a_{t-1} + m\alpha, \qquad b_t = \gamma b_{t-1} + e^{\beta^T u_t} \sum_{j=1}^{m} \lambda_j y_{jt}.$$

共同环境会同时利用所有 m 个路段的信息进行学习，并按其 λ_j 值加权。每个路段对形状参数的贡献为 α ，若有 m 个路段，则每个时间步的有效信息量为 $m\alpha$，从而使后验分布更加集中，可靠性估计也更加精确。

注 2 （可识别性）。 参数 λ_j 和 η_t 仅通过其乘积$\lambda_j\eta_t$进入似然函数：对 $\lambda_j \to c\lambda_j$ 和 $\eta_t \to \eta_t/c$ 进行重新缩放不会改变似然函数。在贝叶斯框架中，*Gamma*先验 $(\eta_0 \mid D^0) \sim$ 伽马分布(a_0, b_0) 为 η_t 锚定尺度； λ_j 则表示各路段的相对 ¯ 率参数。实践中，我们对 λ_j 进行归一化，使其满足 $\lambda = 1$，从而消除尺度歧义（见第7节）。

4.3 联合预测分布

通过对环境进行积分消去，可得到路段行程时间的联合预测分布。为简化记号，我们省略协变量 u_t ，并记为 $\tilde{a} = \gamma a_{t-1}$, $b = \gamma b_{t-1}$。

命题2 (多元复合伽马分布**).** 联合预测密度为

$$p(y_{1t}, \ldots, y_{mt} \mid D^{t-1}) = \frac{\Gamma(m\alpha + \tilde{a})}{\Gamma(\alpha)^m \, \Gamma(\tilde{a})} \, \frac{\prod_{j=1}^{m} \lambda_j^{\alpha} \, y_{jt}^{\alpha-1} \;\cdot\; \tilde{b}^{\tilde{a}}}{\left(\tilde{b} + \sum_{j=1}^{m} \lambda_j y_{jt}\right)^{m\alpha+\tilde{a}}}. \tag{9}$$

证明。 积分 $\int_0^\infty \prod_j p(y_{jt} \mid \eta_t) \cdot p(\eta_t \mid D^{t-1})\, d\eta_t$ 给出伽马密度的乘积，再乘以˜伽马先验分布，并将其作为 Gam$(m\alpha + \tilde{a},\, b + \sum_j \lambda_j y_{jt})$ 的核进行积分。 □

当 $\alpha = 1$（指数观测模型）时，方程（9）简化为动态多元Lomax分布；Nayak（1987）研究了其静态情形，Aktekin 等人（2020）研究了其动态场景。对于一般的 α，它是多元复合伽马分布。

5 路径级行程时间可靠性

本节介绍本文的主要理论贡献：一种闭式路线预测分布，可即时计算可靠性指标。

5.1 通过条件独立性降维

路线行程时间为 $S_t = \sum_{j=1}^{m} y_{jt}$。$S_t$ 的预测分布为

$$P(S_t \le \tau \mid D^{t-1}) = \int_0^\infty P(S_t \le \tau \mid \eta_t)\, p(\eta_t \mid D^{t-1})\, d\eta_t. \tag{10}$$

条件独立结构意味着，$P(S_t \le \tau \mid \eta_t)$涉及独立伽马随机变量之和的分布——这是一个可以高效计算的一维累积分布函数。关于 η_t 的外层积分也是一维积分。

与其他方法相比，计算上的优势显而易见。如果路段相互独立但不服从正态分布，计算 $P(S_t \le \tau)$需要对边缘分布进行 m重卷积（Ng 和 Waller，2010年），并通过快速傅里叶变换进行数值计算。Copula 模型的表现更差：$P(S_t \le \tau)$需要进行 m维数值积分或模拟。相比之下，我们的共同环境表述将问题简化为关于η_t的一维积分，而与 m无关。

5.2 伽马矩匹配近似

给定 η_t，路段行程时间相互独立，且服从 $y_{jt} \mid \eta_t \sim$ 伽马分布$(\alpha, \lambda_j \eta_t)$。条件和 $S_t \mid \eta_t$ 的矩为

$$E[S_t \mid \eta_t] = \frac{\alpha}{\eta_t} \sum_{j=1}^{m} \frac{1}{\lambda_j}, \qquad V[S_t \mid \eta_t] = \frac{\alpha}{\eta_t^2} \sum_{j=1}^{m} \frac{1}{\lambda_j^2}.$$

We a用与这些矩相匹配的单一伽马分布来近似 $S_t \mid \eta_t$：

$$S_t \mid \eta_t \;\approx\; \mathrm{Gam}(\alpha^*,\, c\,\eta_t), \tag{11}$$

其中

$$\alpha^* = \alpha \cdot \frac{\left(\sum_{j=1}^{m} \lambda_j^{-1}\right)^2}{\sum_{j=1}^{m} \lambda_j^{-2}}, \qquad c = \frac{\sum_{j=1}^{m} \lambda_j^{-1}}{\sum_{j=1}^{m} \lambda_j^{-2}}. \tag{12}$$

这种近似有两个重要性质：

- 有效路径形状参数 α^* 不依赖于 η_t。它仅取决于路段参数 λ_j 和每个路段的形状参数 α。

- 有效路径率参数 $c\,\eta_t$ 与 η_t呈线性关系。

这些性质至关重要，因为它们保留了模型的共轭结构，从而能够对 η_t进行闭式积分。

注释 3。当所有 λ_j 都相等（齐次区段）时，$\alpha^* = m\alpha$ 和 $c = \lambda$，因此该近似是精确的：独立同分布的伽马分布$(\alpha, \lambda\eta_t)$ 随机变量之和恰好服从伽马分布$(m\alpha, \lambda\eta_t)$。更一般地，

已知，当各个伽马随机变量的形状参数适中时，独立伽马变量之和的伽马近似是准确的（*Moschopoulos，1985*年）。

5.3 路径预测分布的闭式解

将矩匹配近似（11）代入路线积分（10），得到与单变量情形相同的复合伽马结构。

命题3（路径预测分布）。 在伽马分布矩匹配近似下，路线行程时间的预测分布为

$$p(S_t \mid D^{t-1}) = \frac{\Gamma(\alpha^* + \tilde{a})}{\Gamma(\alpha^*)\,\Gamma(\tilde{a})} \, \frac{c^{\alpha^*}\, S_t^{\alpha^*-1}\, \tilde{b}^{\tilde{a}}}{(c\,S_t + \tilde{b})^{\alpha^*+\tilde{a}}}, \tag{13}$$

其中 $\tilde{a} = \gamma a_{t-1}$ *a* ˜*nd* $b = \gamma b_{t-1}$ 是多变量滤波器的先验参数。等价地 ，

$$\frac{\tilde{a}}{\alpha^*} \cdot \frac{c\,S_t}{\tilde{b}} \;\sim\; F(2\alpha^*,\, 2\tilde{a}). \tag{14}$$

证明。 积分 $\int_0^\infty \mathrm{Gam}(S_t;\, \alpha^*, c\eta_t) \cdot \mathrm{Gam}(\eta_t;\, \tilde{a}, \tilde{b})\; d\eta_t$ 根据标准共轭性计算得到复合伽马分布（13）。 F分布关系则源自贝塔素分布与 F分布之间众所周知的联系。

□

这一结果是本文的核心贡献。这意味着：

- 路线预测累积分布函数为

$$P(S_t \le \tau \mid D^{t-1}) = F_{2\alpha^*,\,2\tilde{a}}\!\left(\frac{\tilde{a}}{\alpha^*} \cdot \frac{c\,\tau}{\tilde{b}} \right), \tag{15}$$

  其中 $F_{\nu_1,\;\nu_2}$ 表示自由度为 ν_1 和 ν_2的 F 分布的累积分布函数。

- 水平 q 处的路线预测分位数为

$$S_t^{(q)} = \frac{\alpha^*\,\tilde{b}}{c\,\tilde{a}} \cdot F^{-1}_{2\alpha^*,\,2\tilde{a}}(q). \tag{16}$$

- 通过标准统计函数，两者均可在 $O(1)$ 的计算时间内得到。

5.4 路线可靠性指标

利用闭式路径预测分布，可以在每个时间步即时计算联邦公路管理局（2006）和 Lomax 等人（2003）推荐的标准可靠性指标。

准时到达概率。 对于阈值 τ（例如，自由流行程时间加上缓冲时间），

$$P_{\text{on-time}}(t) = P(S_t \le \tau \mid D^{t-1}) = F_{2\alpha^*,2\tilde{a}}\!\left(\frac{\tilde{a}\,c\,\tau}{\alpha^*\,\tilde{b}} \right).$$

规划时间指数（PTI）。 行程时间第95百分位数与自由流行程时间之比：

$$\mathrm{PTI}(t) = \frac{S_t^{(0.95)}}{S_{\mathrm{ff}}},$$

其中 S_{ff} 为自由流路线行驶时间， $S_t^{(0.95)}$ 由式（16）在 $q = 0.95$下给出 .

缓冲 指数。 第95百分位数与中位数之间的归一化差值 n:

$$\mathrm{BI}(t) = \frac{S_t^{(0.95)} - S_t^{(0.50)}}{S_t^{(0.50)}}.$$

这三项指标均为在当前先验分布param-˜eters（$\tilde{a}$， b）处求值的 F分布分位数的函数，因而可以实时计算。这与现有方法形成鲜明对比：现有方法中的可靠性指标要么根据历史样本计算（静态，非实时），要么需要模拟（Taylor，2013年）。

6 贝叶斯推断

6.1 通过吉布斯采样和前向滤波后向抽样进行马尔可夫链蒙特卡罗

为了从完整数据集 D^T 进行批量推断，我们构建了一个 Gibbs 采样器。所有未知参数 $\Psi = (\lambda,\ \alpha,\ \gamma,\ \beta)$ 与环境过程 $\eta^T = (\eta_1,\ \ldots\ldots,\ \eta_T)$ 的联合后验分布无法以封闭形式求得，但模型结构使高效的分块抽样成为可能。

Gibbs 采样器在以下步骤之间交替进行：

1. **抽取 η^T，给定 Ψ 和 D^T**：我们采用 Fruhwirth-Schnatter（1994）的前向滤波后向抽样（FFBS）算法。前向步骤通过多元递推 $a_t = \gamma a_{t-1} + m\alpha$,$b_t = \gamma b_{t-1} + \sum_j \lambda_j y_{jt} e^{\beta^T u_t}$ 计算滤波分布 $(\eta_t \mid \Psi,\ D^t) \sim$伽马分布$(a_t,\ b_t)$（单变量情形参见公式 (3)）。后向步骤抽取 $\eta_T \sim$伽马分布$(a_T,\ b_T)$，然后对于 $t = T-1,\ \ldots\ldots,\ 1$，使用 Aktekin 等人（2020）的移位伽马分布表征：抽取 $z_t \sim$伽马分布$((1-\gamma)a_t,\ b_t)$，并令 $\eta_t = \gamma\eta_{t+1} + z_t$。这种方法无需基于截断的抽样，且易于高效实现。

2. **抽样 λ_j给定 η^T、其他参数和 D^T**：采用独立的伽马分布先验$\lambda_j \sim$伽马分布$(r_{j0},\ q_{j0})$，其完全条件分布为

$$(\lambda_j \mid \eta^T, D^T, \alpha, \beta) \sim \mathrm{Gam}\left(r_{j0} + T\alpha,\ q_{j0} + \sum_{t=1}^{T} \eta_t y_{jt} e^{\beta^T u_t}\right).$$

λ_j 在给定 η^T 的条件下相互独立。

3. **抽样 γ**：折扣因子没有共轭的全条件分布。遵循 Aktekin 等人（2020）的做法，我们采用 Metropolis 步骤，以边际后验分布 $p(\gamma \mid$

$D^T,\ \alpha,\ \beta,\ \lambda) \propto \prod_t p(\mathbf{y}_t \mid D^{t-1},\ \gamma,\ \alpha,\ \beta,\ \lambda) \cdot p(\gamma)$ 为目标，其中乘积项是前向滤波器给出的单步预测密度，且已对 η 积分消去。以当前值为中心的 Beta 提议分布可提供一种自然的有界提议（Gamerman 和 Lopes，2006）。

4. **抽样 α 与 β**：这些参数也通过 Metropolis 步骤更新，并使用对 η 积分消去后的相同边际似然。

当 (α, γ) 固定时——例如，通过第 7 节的网格搜索进行经验贝叶斯选择——步骤 (iii) 和 (iv) 就不再需要，采样器简化为交替进行针对 η^T的前向滤波后向抽样，以及针对 λ_j 的共轭 Gamma 抽样；两者都是精确的（无需 Metropolis 步骤）。这种条件吉布斯采样器 由于具有完全共轭结构而能快速混合，并可得出潜在环境过程和路段速率参数的后验分布，从而量化点估计之外的参数不确定性。

6.2 用于实时更新的粒子滤波

对于实时更新，Gibbs 采样器并不实用，因为每次有新观测到达时都需要重新处理所有数据。当 $(\alpha, \gamma, \lambda)$ 固定时（例如，按照第 7 节所述通过网格搜索选定），命题 1 中的共轭更新可为 η_t 提供精确的在线推断，无需近似。当 Ψ也必须在线估计时，我们使用粒子学习方法（Carvalho 等人，2010年），从时间 $t-1$ 递归更新 (η_t, Ψ) 的后验分布至时间 t。

在每个时间步，加权粒子集合 $\{(\eta_t^{(i)},\ \Psi^{(i)},\ w_t^{(i)})\}_{i=1}^N$ 表示后验分布。当新数据 $(y_{1t},\ \ldots,\ y_{mt},\ u_t)$到达时：

1. 推进每个粒子的环境：$\eta_t^{(i)} \sim$ 伽马分布$(\gamma a_{t-1}^{(i)},\ \gamma b_{t-1}^{(i)})$。

2. 使用观测似然更新权重。

3. 如果有效样本量低于阈值，则进行重采样。

4. 更新 λ_j 的充分统计量：由于 $(\lambda_j \mid \eta^t,\ D^t) \sim$服从伽马分布$(r_{jt},\ q_{jt})$，其中参数为 $r_{jt} = r_{j,\ t-1} + \alpha$ 和$q_{jt} = q_{j,\ t-1} + \eta_t y_{jt} e^{\beta^T u_t}$，因此更新只需递增充分统计量。

粒子滤波器可提供实时后验分布和预测密度，因此适合在交通管理系统中实际部署（Polson and Sokolov，2015；Mihaylova 等，2007）。

7 应用：I-55 史蒂文森高速公路

7.1 数据描述

我们将该模型应用于伊利诺伊州芝加哥市史蒂文森高速公路（I-55）北向路段的交通数据。这些数据来自盖瑞-芝加哥-密尔沃基（GCM）走廊交通管理系统，包含环形线圈检测站每小时记录的速度、流量和占有率。

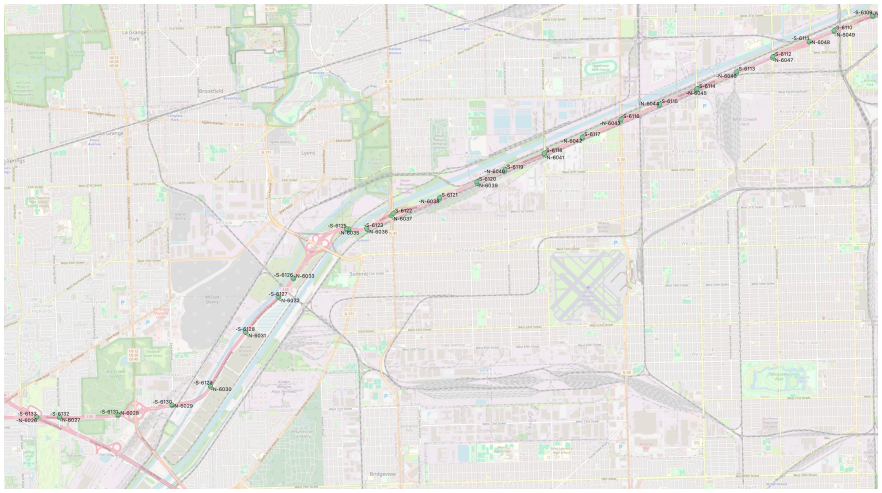

图 1：I-55州际公路北向路段（史蒂文森高速公路）上16个环形线圈检测站的位置，全长8.26英里，从I-294州际公路附近的传感器 6030延伸至戴门大道附近的传感器 6045。

我们的研究走廊由连续的 16 个环形线圈检测站（传感器 6030–6045）组成，覆盖史蒂文森高速公路北向路段的 8.26 英里范围（图 1）。传感器 6034 和 6035 安装在同一地点（相距 31 米），因此形成了一段长度近乎为零的路段；由于两者各自报告独立的速度测量值，故均予以保留。表 2 汇总了根据传感器坐标利用哈弗辛公式计算得出的路段距离和平均行程时间。

我们关注2019年（数据最完整的一年）周三下午的情况，每天分为七个一小时时段（下午2:00–2:59至晚上8:00–8:59）。这提供了 $52 \times 7 = 364$个潜在时间段；剔除任何传感器存在数据缺失的时段后，剩下248个完整观测值。

路段级行程时间计算为 $y_{jt} = d_j / v_{jt}$，其中 d_j 是从传感器 j到下游相邻传感器的距离（英里），v_{jt} 是传感器 j 的小时平均速度（英里/小时），由此得到以小时为单位的行程时间，再将其换算为分钟。对于最后一个传感器（6045），距离设为与前一个路段相同。路线行程时间为 $S_t = \sum_{j=1}^{m} y_{jt}$，其中 $m = 16$。

7.2 静态分布分析

图 2 展示了四个代表性传感器在路段层面的分布异质性。每个传感器的形状各不相同，从传感器6030（$\hat{}\alpha = 34.2$，路段较短、速度适中）近乎对称的分布，到传感器6037（$\hat{}\alpha = 4.1$，反复出现的瓶颈区域）明显右偏的分布。伽马分布能够很好地拟合这一完整的形状范围，而正态分布则无法体现正支撑和偏度。

路线行程时间的均值为 28.8 分钟，标准差为 8.2 分钟，范围为 16.9 至 52.8 分钟（图3）。表 3 展示了几种静态分布的拟合优度。伽马分布的柯尔莫戈洛夫 – 斯米尔诺夫检验 p- 值最高，而逆高斯分布的赤池信息准则和贝叶斯信息准则最低；这三种偏斜分布的表现都显著优于对称的正态分布，后者无法

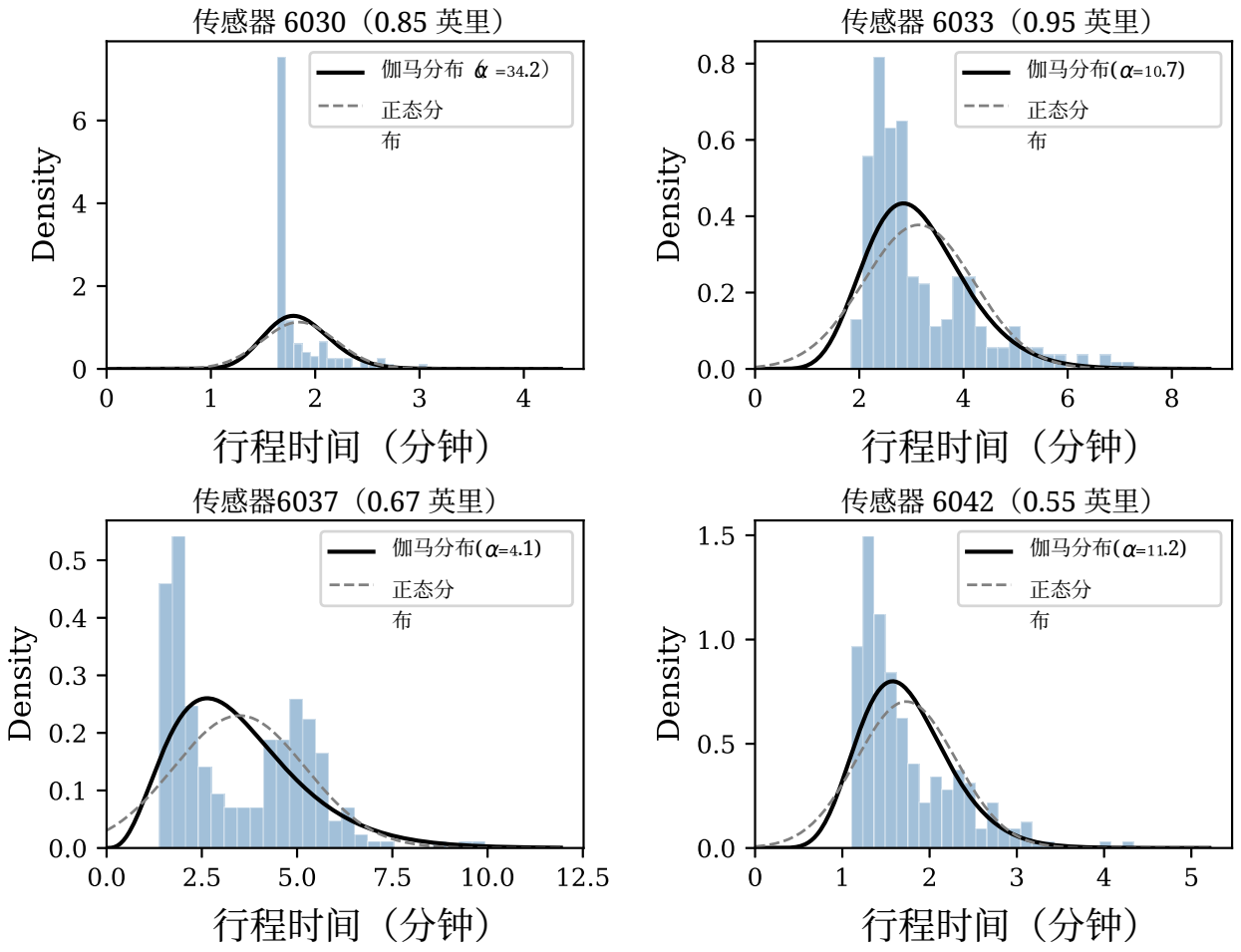

图 2：四个代表性传感器的路段级旅行时间分布，变异性从低（传感器 6030，$\hat{\alpha} = 34.2$）到高（传感器6037，$\hat{\alpha} = 4.1$）。伽马分布（实线）捕捉了数据中存在的右偏，而正态分布（虚线）无法适应正支撑和非对称性。每个路段呈现出不同的分布特征，从而促使采用路段特定的 λ_j 参数。

ˆ 表 2：路段距离和平均行驶时间（2019 年星期三下午 2 点至 8 点）。 λ 数值经过四舍五入；表 5 中的 $\alpha^* = 13.3$ 根据全精度值计算得出。

传感器	距离（英里）	平均速度（英里/小时）	平均行驶时间（分钟）	伽马分布 ˆα	$\hat{\lambda}$
6030	0.853	28.5	1.84	34.2	0.46
6031	0.632	17.1	2.44	9.3	0.35
6032	0.313	21.2	0.96	11.1	0.88
6033	0.949	19.8	3.14	10.7	0.27
6034	0.031	20.4	0.10	7.9	8.34
6035	0.239	16.6	1.04	4.9	0.82
6036	0.370	20.3	1.20	11.0	0.71
6037	0.666	14.7	3.49	4.1	0.24
6038	0.536	15.0	2.47	7.5	0.34
6039	0.400	22.0	1.16	16.4	0.73
6040	0.565	19.2	1.83	26.1	0.47
6041	0.539	19.4	1.93	6.3	0.44
6042	0.555	20.9	1.73	11.2	0.49
6043	0.547	20.5	1.92	5.3	0.44
6044	0.530	16.6	2.09	12.6	0.41
6045	0.530	23.3	1.46	14.7	0.59
总计	8.26		28.80		

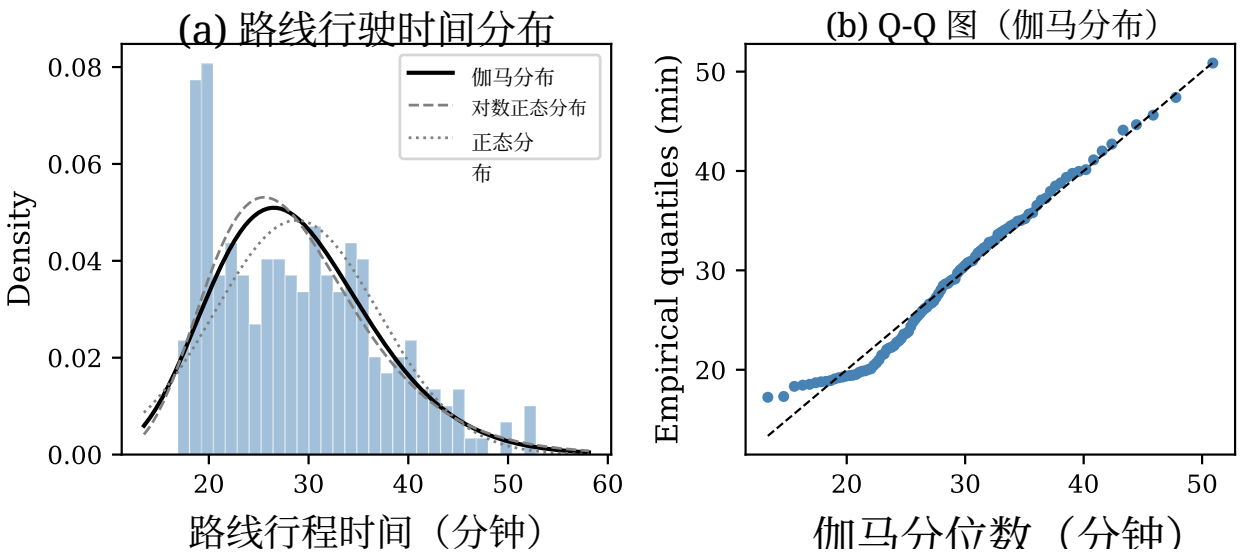

图3：(a) 使用伽马分布、对数正态分布和正态分布拟合的路线级行程时间分布。(b) 经验分位数与拟合的伽马分布的 Q-Q 图。

使其支撑集为（$-\infty, \infty$），这与 Susilawati et al. (2013) 和 Zang et al. (2022b) 的研究结果一致。

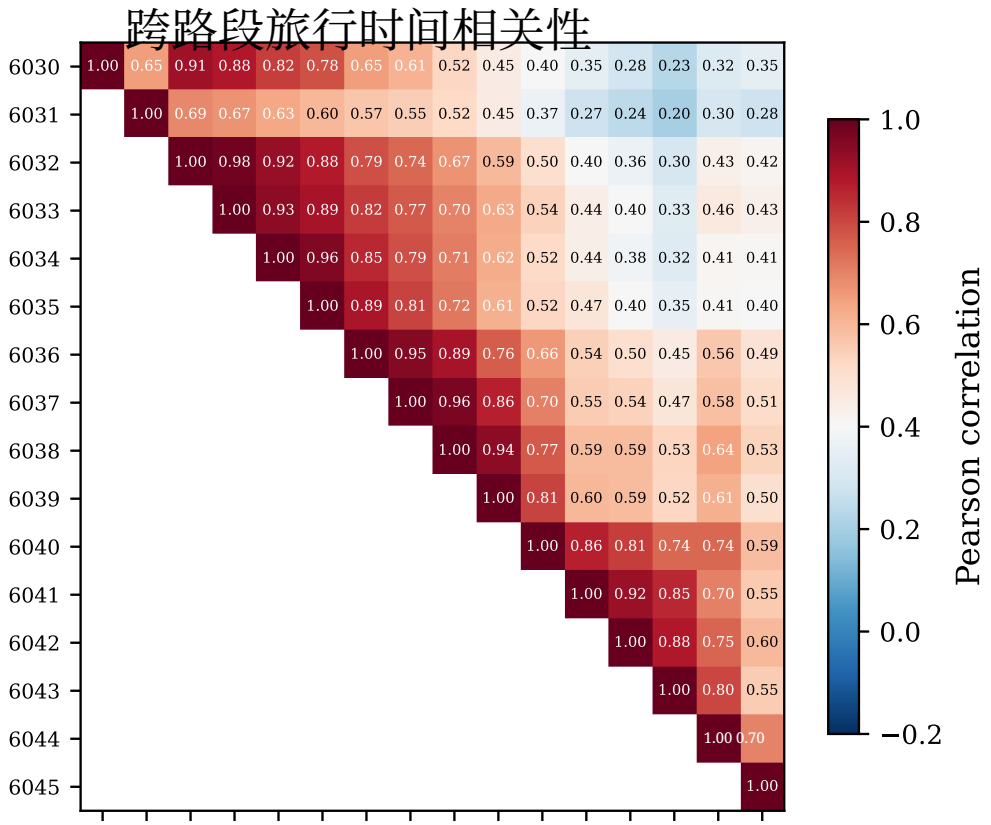

图4：跨路段旅行时间相关矩阵。相邻路段呈现较强的正相关（0.6–0.8），即使相距较远的路段之间仍保持正相关，这支持共同随机环境假设。

一个重要的经验观察是，路段行程时间之间存在较强的正相关性（图4）。相邻路段的相关性为 0.6–0.8，即使相距较远的路段之间，相关性也达到 0.2–0.4。这种普遍存在的正向依赖性正是共同随机环境 η_t 旨在捕捉的：共同的交通状况（拥堵波、天气、需求冲击）会同时影响整个走廊。

双峰性与状态结构。 虽然单成分伽马分布能够充分拟合路径分布（表3），但它掩盖了一个重要特征。图5显示了包含 $K = 1$、2、3 个部件的伽马混合拟合。双成分混合分布（$K = 2$，贝叶斯信息准则 $= 1698$，而$K = 1$为1739）揭示出清晰的双峰结构：一个以 $\mu = 19$ 分钟为中心的窄自由流分量（权重为0.16），以及一个以 $\mu = 31$ 分钟为中心的较宽拥堵分量（权重为0.84）。第三个部件（$K = 3$，贝叶斯信息准则 $= 1710$）并未改善拟合效果。这种双峰性反映了

表3：路线行程时间的静态分布拟合（248个观测值）。

分布	KS统计量	KS p值	AIC	BIC
伽马分布	0.080	0.077	1732.0	1739.0
对数正态分布	0.085	0.055	1727.7	1734.7
逆高斯分布	0.084	0.055	1726.0	1733.0
威布尔	0.092	0.029	1758.3	1765.4

在观测期间不同交通状态——自由流状态与拥堵状态——之间的交替。单一静态分布会对这些状态进行平均；第3节提出的动态模型则直接处理了这一问题，因为随时间变化的环境 η_t 会在路况良好时将预测分布推向自由流模式，并在路况恶化时将其推向拥堵模式。

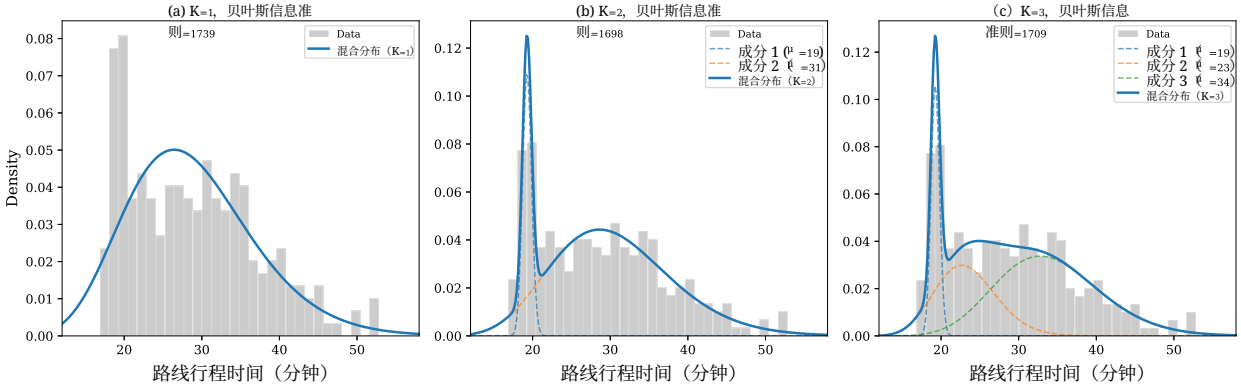

图5：路线行程时间的伽马混合分解。(a) 单一伽马分布（$K = 1$）。(b) 双成分混合分布（$K = 2$，贝叶斯信息准则优选）：一个较窄的自由流分量（$\mu = 19$ 分钟）和一个较宽的拥堵分量（$\mu = 31$ 分钟）。(c) 三个成分（$K = 3$）并未改善贝叶斯信息准则。双峰性反映了自由流状态与拥堵状态之间的状态切换。

分段伽马拟合显示，形状参数 $\hat{\alpha}$存在显著异质性，范围从 4.1（传感器6037，变异性高）到 34.2（传感器 6030，变异性低）。我们的多变量模型通过路段均值捕捉路段特定的λ_j 参数所体现的这种异质性，而共同环境 η_t 则通过共享条件将所有路段联系起来。不过需要指出的是，多变量模型强制所有路段共用一个共同形状参数 α；只有率参数 $\lambda_j \eta_t$ 因路段而异。这是保留共轭结构所付出的代价，而该结构使路线分布能够得到闭式表达。拟合效果最佳的共同 $\alpha = 1$ （第7节）与各路段的最大似然估计相差甚远，因此只能近似捕捉路段级伽马形状参数。尽管存在这种权衡，由于对 m 路段进行矩匹配聚合后得到的有效路线形状参数 α^* （公式12）适用于路径级预测，因此路径级校准仍然非常出色，即使共同 α 未能与每个路段完全匹配。

7.3 经验贝叶斯选择(α, γ)

在此应用中，我们将协变量向量设为 $u_t = 0$ （等价地， $\beta = 0$），因此所有时变条件都由潜在环境 η_t吸收。原则上，通过 $e^{\beta^T u_t}$ 项纳入观测协变量（天气、事故）很直接，但小时级环形线圈检测器数据不包含同期协变量测量值。

我们的推断分为两个阶段。首先，我们通过经验贝叶斯方法选择超参数（α, γ），并在一个网格上评估路径级预测校准。其次，在给定所选（α, γ）的条件下，我们运行条件吉布斯采样器（第 6 节），以获得路段速率参数 λ_j 和潜在环境过程 η_t的完整后验分布。采用两阶段方法的原因是，（$\alpha, \gamma, \lambda, \eta^T$）的完整联合后验会集中于一个使路段级边际似然最大化、但路径级预测效果不佳的众数；经验贝叶斯步骤则选择（α, γ），以优化路径级校准。

在网格搜索中，我们将 λ_j 初始化为与 $1/\bar{y}_j$ 成比例（即路段样本均值的倒数），并将其标准化为均值为 1，从而消除第 4 节讨论的尺度歧义。我们评估两种方法：一种是直接应用于路线行程时间的单变量模型，另一种是应用于各路段行程时间的多变量模型，并通过矩匹配公式（14）计算路径可靠性。

为评估模型，我们采用三个标准：

1. **对数预测似然：** $\sum_{t>t_0} \log p(y_t \mid D^{t-1})$，预热期 $t_0 = 30$之后单步预测密度对数之和。

2. **PIT校准：** 如果模型校准良好，概率积分变换值 $u_t = P(Y_t \le y_t \mid D^{t-1})$应服从 Uniform(0,1)分布。我们采用均匀性的柯尔莫哥洛夫-斯米尔诺夫检验。

3. **覆 盖率：** 落在90%预测区间内的观测值比例 al.

单变量路径模型。表 4 展示了对 α 和 γ进行网格搜索的结果。采用$\alpha = 10$ 的模型在一系列折扣因子下都能提供良好校准的预测，其90%区间覆盖率始终接近名义水平。

表4：路线行程时间的单变量动态伽马模型：选定的参数组合。

α	γ	对数预测似然	PIT KS统计量	柯尔莫戈洛夫–斯米尔诺夫检验 p-值	90%覆盖率
10	0.80	−774.2	0.082	0.103	0.977
10	0.85	−771.7	0.071	0.211	0.977
10	0.90	−769.7	0.075	0.165	0.972
10	0.95	−768.0	0.074	0.171	0.968
10	0.99	−768.2	0.068	0.247	0.963

包含路径可靠性的多变量模型。该多变量模型应用于由路段行程时间构成的 m维向量，并通过闭式F分布公式（14）计算路径可靠性。图6展示了在（α、 γ）网格上的路径级PIT校准（KS p值）。在 $\alpha \in [1.0, 1.5]$ 和 $\gamma \in [0.50, 0.70]$处出现了明显的良好校准区域，其中 $\alpha = 1.0$、 $\gamma = 0.70$ 处的校准效果最佳（KS $p = 0.467$）。表 5 展示了校准效果最佳的组合的详细结果。这些数值将用于后续的贝叶斯后验分析（第7.4节）。

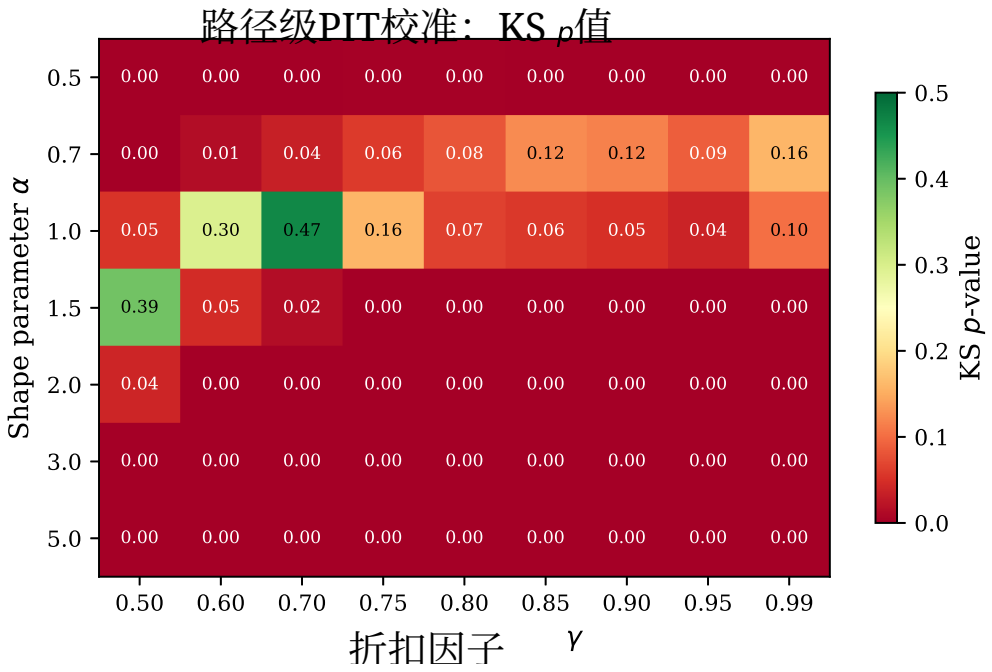

图6：(α, γ) 的经验贝叶斯选择：超参数网格上的路径级PIT校准（KS p值）。绿色单元格表示校准良好的模型（$p > 0.05$）。所选值$\alpha = 1.0$, $\gamma = 0.70$ ($p = 0.467$)用于后续的贝叶斯后验推断。

表 5：多元动态伽马分布模型：通过矩匹配进行路径级校准。

α	γ	联合对数似然	路径KS检验	KS p	路径90%覆盖率	α^*
1.00	0.70	−4927	0.057	0.467	0.954	13.3
1.50	0.50	−4203	0.060	0.389	0.931	19.9
1.00	0.60	−4930	0.065	0.295	0.972	13.3
1.00	0.65	−4928	0.067	0.269	0.963	13.3
0.70	0.99	−5603	0.075	0.159	0.963	9.3

采用 $\alpha = 1.0$ （对应于给定η_t的指数分布路段行程时间）和 $\gamma = 0.70$ 的多变量模型实现了最佳路径校准：概率积分变换KS检验给出的名义 p值为0.47（不过请参见第7节中的自相关注意事项），90%区间覆盖率为95.4%。有效路径形状参数 $\alpha^* = 13.3$ （使用最大似然估计 λ_j）提供了形状良好的预测

分布 。我们在第7.4节的完整贝叶斯后验分析中采用这些数值 .

排名第二的 $\alpha = 1.5, \gamma = 0.50$ 也实现了良好的校准（KS $p = 0.389$），且环境过程的适应速度更快。选择 $\gamma = 0.50$–0.75意味着模型对近期观测值赋予了较大权重：对于小时数据， $\gamma = 0.70$ 对应的有效记忆大约为 $1/(1-\gamma) \approx 3.3$ 小时，这意味着下午较早时段的状况到了傍晚时会被大幅淡化。这种较短的记忆与高峰时段城市快速路上快速变化的状况相符，因为拥堵可能在1–2小时内形成并消退。

一个重要发现是，最佳单变量模型的 $\alpha = 10$ （路线行程时间的伽马形状参数适中且峰度较高），而最佳多变量模型在路段层级的 $\alpha = 1$–1.5。路线层级的形状参数 $\alpha^* = 13$–20是对 $m = 16$ 个路段进行聚合后得到的，这表明矩匹配公式（12）能够恰当地放大有效形状参数。值得注意的是， $\alpha = 1$ 对应于给定 η_t时服从指数分布的路段行程时间——这是伽马分布族中变异性最大的成员。已知，对 m个独立指数分布随机变量之和采用伽马矩匹配近似，即使在 m适中的情况下也很准确（Moschopoulos，1985）；而有了 $m = 16$ ，所得的 $\alpha^* \approx 13$ 便能形成形状良好的路线分布。

7.4 贝叶斯后验分析

在通过经验贝叶斯（第7.3节）选定 $(\alpha,\ \gamma) = (1。0,\ 0。70)$ 后，我们运行第6节中的条件吉布斯采样器，以获得路段速率参数 λ_j 和潜在环境过程 $\eta^T = (\eta_1,\ \ldots\ldots,\ \eta_T)$ 的完整后验分布。每次迭代包含两个共轭步骤：(i) 对 η^T进行前向滤波和 FFBS 后向抽样（Fruhwirth-Schnatter，1994；Aktekin et al.，2020）；(ii) 根据 η^T，从各自的伽马完全条件分布中抽取每个 λ_j。我们从不同的初始值运行4条链，每条链迭代10,000次（预热期2,000次，每隔1次抽取一次），共得到16,000个后验样本。完全共轭结构带来了出色的混合性：所有 λ_j 中的最大分裂-R为1。0004，有效样本量最小值为3，722。

表 6 报告了 λ_j的后验汇总。后验均值接近表 2 中的最大似然估计值，且95%可信区间较窄，表明数据对这些参数具有较强的信息量。路径层面的形状参数后验为 $\alpha^* = 13.26$ ，其95%可信区间为$[12.91, 13.57]$，与13.3的代入值一致。

图7显示了 η_t 的后验轨迹及逐点95%可信区间。环境过程与观测到的拥堵呈反向变化：在自由流状态下（行程时间较短）上升，在拥堵时段下降。相对于后验均值，可信区间较窄，表明滤波后验能够在每个时间步确定环境状态，且不确定性有限。

图8显示了 λ_j 的后验分布及最大似然估计值。对于全部16个路段，后验分布都紧密集中在最大似然估计值附近；其中传感器6037的不确定性相对最大（该路段的平均行程时间最高且 λ_j最低）。较大的 $\lambda_{6034} \approx 8.4$反映了共址传感器对6034/6035之间的距离接近于零。

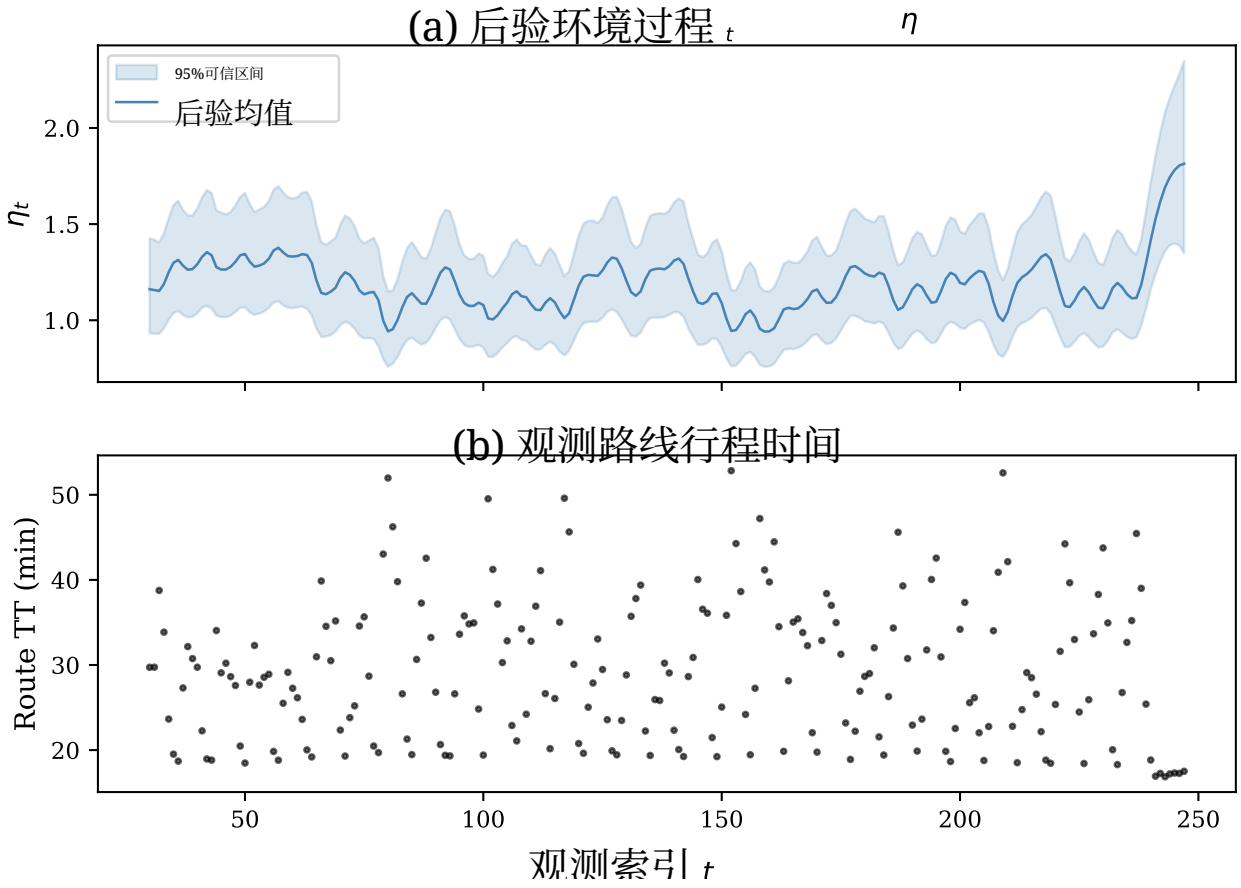

图7：(a) 条件吉布斯采样器得出的共同环境过程 η_t 的后验均值和95%可信区间。(b) 作为参考的观测路线行程时间。环境过程与拥堵程度呈负相关：自由流时段 η_t 会上升，而拥堵时段则会下降。

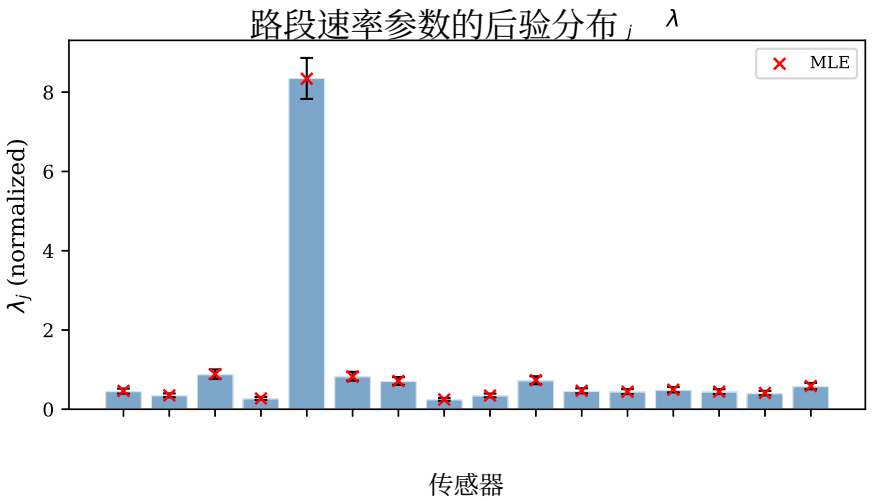

图8：路段速率参数的后验分布 λ_j：柱形表示后验均值及95%可信区间（误差棒），叉号表示最大似然估计值。所有路段的后验分布都紧密集中在最大似然估计值附近。

表 6：条件吉布斯采样器 ($\alpha = 1.0$, $\gamma = 0.70$) 得到的路段速率参数 λ_j 后验汇总。后验均值、95%可信区间和最大似然估计值。

传感器	后验均值	95%可信区间	MLE
6030	0.456	$[0.395, 0.520]$	0.462
6031	0.349	$[0.302, 0.399]$	0.350
6032	0.882	$[0.769, 1.005]$	0.884
6033	0.272	$[0.235, 0.311]$	0.272
6034	8.359	$[7.830, 8.864]$	8.338
6035	0.828	$[0.722, 0.944]$	0.820
6036	0.711	$[0.618, 0.812]$	0.713
6037	0.247	$[0.213, 0.284]$	0.245
6038	0.345	$[0.300, 0.396]$	0.345
6039	0.728	$[0.632, 0.831]$	0.734
6040	0.461	$[0.401, 0.528]$	0.466
6041	0.444	$[0.383, 0.507]$	0.443
6042	0.490	$[0.424, 0.560]$	0.492
6043	0.442	$[0.383, 0.504]$	0.444
6044	0.405	$[0.352, 0.464]$	0.409
6045	0.581	$[0.504, 0.664]$	0.586

对 $(\lambda,\ \eta^T)$ 的后验样本取平均后，贝叶斯预测分布的路径级柯尔莫戈洛夫–斯米尔诺夫检验 p-值为 0.41，90%覆盖率为 95.4%——与代入估计结果几乎完全相同。这证实，对于此应用而言，数据提供的信息足够充分，使代入估计和贝叶斯预测分布实际上等效。实际意义在于，第 5 节中的闭式共轭更新可直接用于实时可靠性计算，与完整的贝叶斯推断相比，损失可以忽略不计。

7.5 路径可靠性结果

使用参数为 $(\alpha,\ \gamma) = (1.0,\ 0.70)$ 的多变量模型和最大似然估计 λ_j（其结果与后验均值几乎完全相同；见表 6），我们通过命题 1 的共轭在线更新，在每个时间步计算路径可靠性指标；这是实际应用中相关的模式。对于这条长 8.26 英里的走廊，自由流路线行驶时间（第 5 百分位数）约为 18 分钟，相当于平均速度 27.5 英里/小时。

图 9 展示了观测路线行程时间，以及叠加其上的多变量模型一步超前预测区间。预测带会随条件变化而调整：在高拥堵时段变宽（反映出更大的不确定性），在自由流时段变窄。几乎所有观测值都落在 90% 预测区间内（实际覆盖率：95.4%），且中位数很好地反映了中心趋势。

图 10 给出了正式的校准评估。PIT 直方图近似均匀，PIT 经验累积分布函数完全位于 95% 柯尔莫哥洛夫–斯米尔诺夫置信带内，其中 $p = 0.47$。我们注意到，90% 预测区间的覆盖率达到 95.4%，高于名义上的 90%，这表明存在轻微的

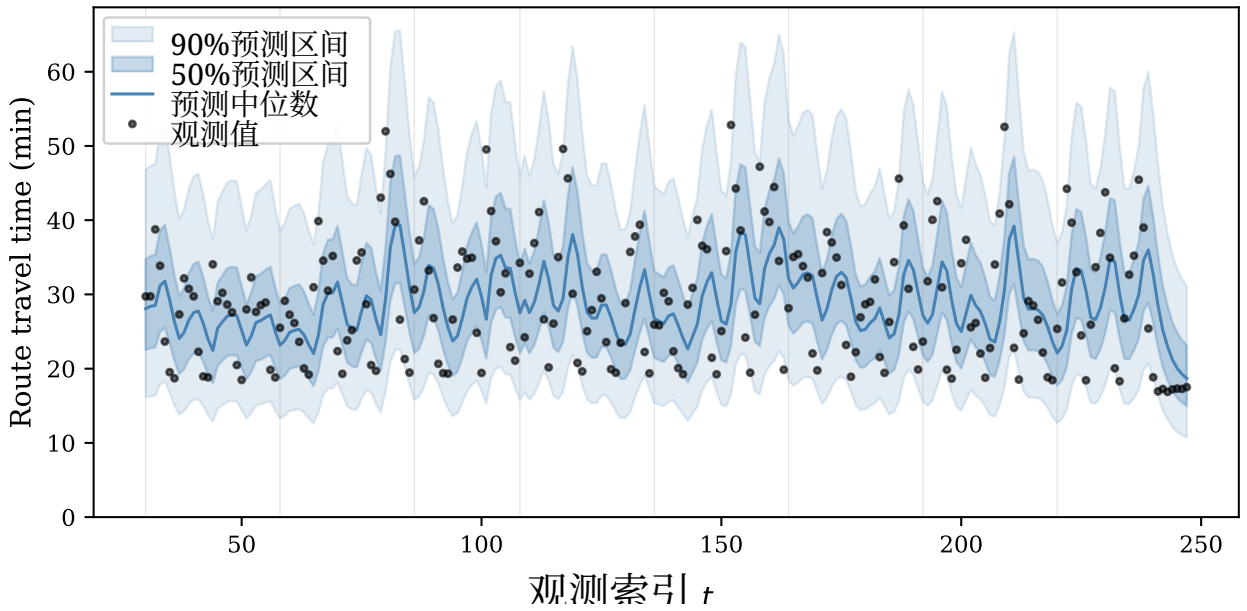

图 9：观测路线行程时间（圆点）和多元动态伽马模型的一步超前预测区间（$\alpha = 1.0$, $\gamma = 0.70$）。50%（深色）和90%（浅色）预测区间随条件变化而调整：拥堵时段区间变宽，自由流时段区间变窄。

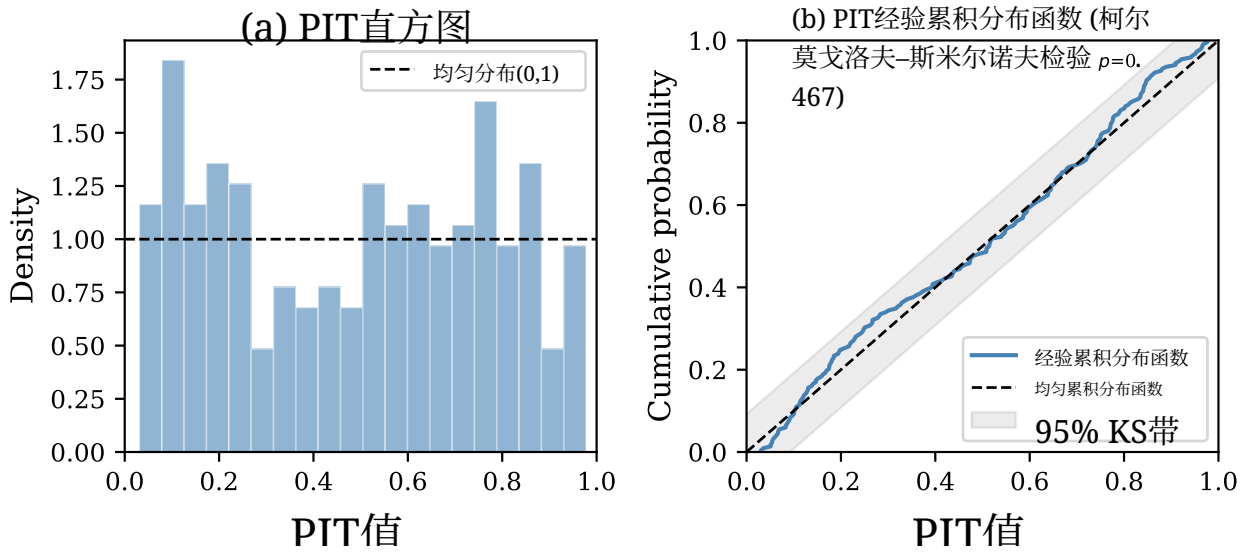

图 10：路径级预测分布的概率积分变换（PIT）校准。(a) PIT直方图：校准良好的模型会产生均匀分布的直方图。(b) PIT 值的经验累积分布函数与均匀分布(0,1)累积分布函数的对比，灰色区域为95%柯尔莫哥洛夫–斯米尔诺夫置信带。

保守（过度离散）的预测尾部。这种过度覆盖与矩匹配近似一致，该近似可能会略微高估路径级分布的尾部权重。KS检验在 5% 显著性水平下未能拒绝均匀性，但在经过预热期后，只有 $n = 218$ 个评估点，因此该检验检测中等程度偏离的能力有限。我们还注意到，标准 KS 检验假设 PIT 值独立同分布，而预测的序贯结构会使 PIT 值产生显著的自相关（滞后 1 阶自相关为 0.52；Ljung–Box检验 $Q = 59.8$，自由度为 1，$p < 10^{-14}$）。这种自相关自然源于日内时间结构：同一个周三下午相邻时段的交通状况相似。虽然 PIT 值的边缘分布看起来近似均匀，但序列相依意味着 KS p值高估了有效样本量，因而夸大了校准证据。采用考虑时间相依的更复杂校准检验（例如 Diebold 等人，1998）将能提供更严谨的评估。

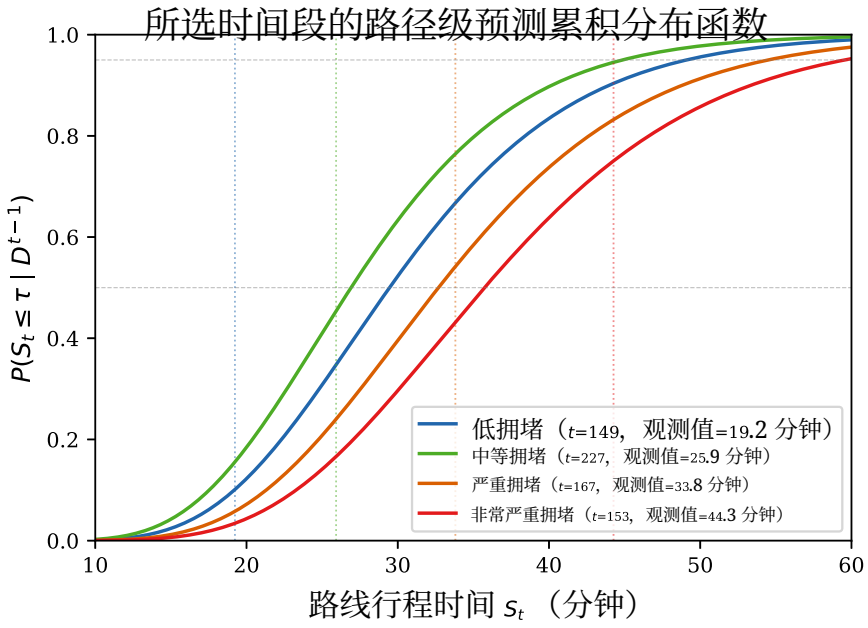

图11：代表不同拥堵程度的四个选定时间段的路径级预测累积分布函数。在拥堵情况下，闭式 F 分布（命题3）向右移动，且分布更加分散。

图11展示了覆盖不同拥堵程度的四个时间段的路径预测累积分布函数。在低拥堵时，分布集中在20分钟左右，累积分布函数陡峭；在严重拥堵时，分布向右移至 40+分钟，累积分布函数则更平缓，反映出更大的不确定性。所有四个累积分布函数均通过 F 分布公式（14）即时计算得出。

图12展示了根据模型的闭式预测分布实时计算的三项可靠性指标。规划时间指数（PTI）从轻度交通时的2.0到高峰拥堵时的3.5不等，这意味着为了确保准时到达概率达到95%，出行者必须预留相当于自由流行程时间2至3.5倍的时间。准时到达概率（面板c）在不同时间段之间差异显著，在最严重的拥堵时段会降至30%以下。这些

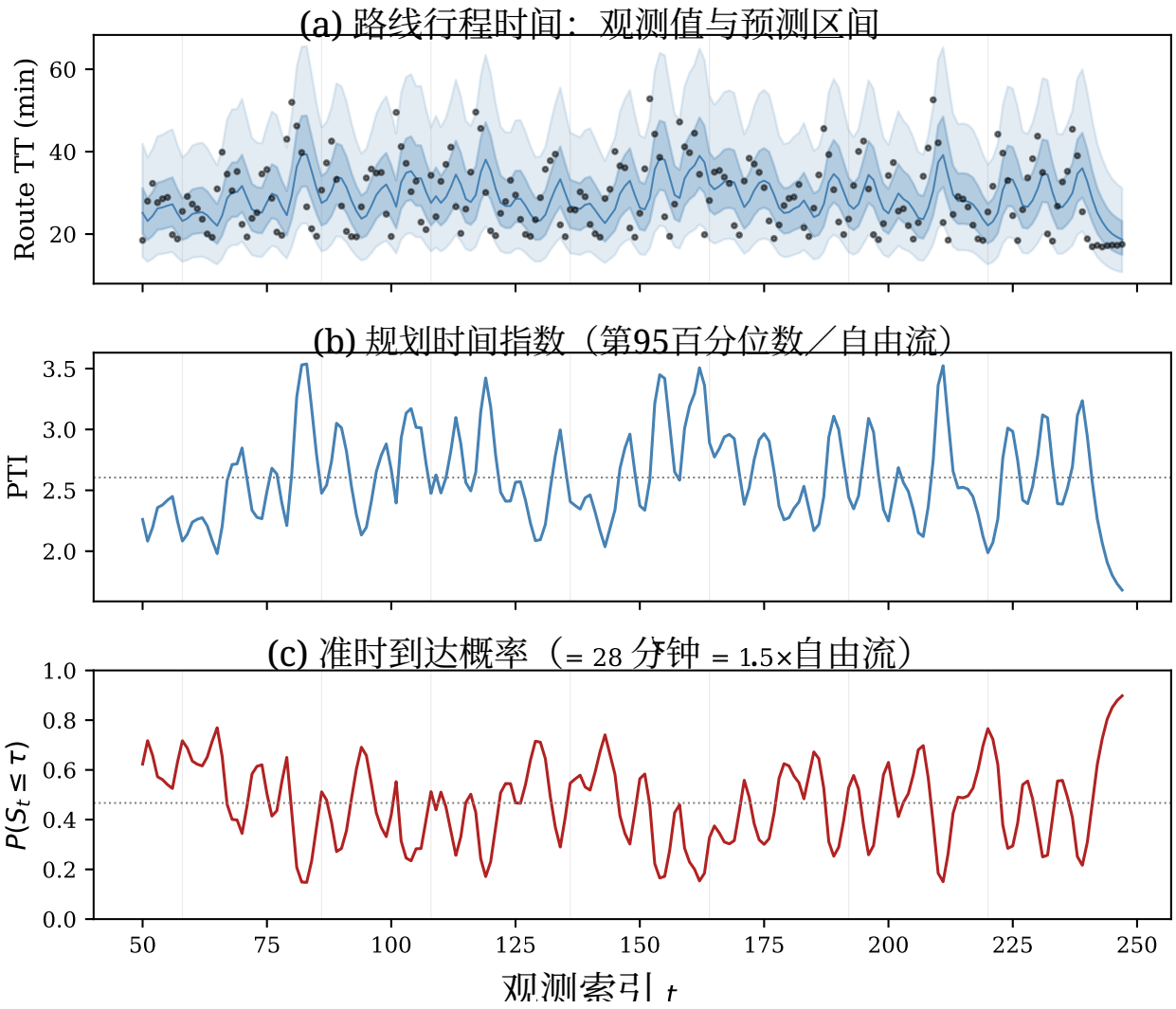

图12：根据多变量模型计算的动态可靠性指标。(a) 带预测区间的路线行程时间。(b) 规划时间指数（第95百分位数除以自由流行程时间）。(c) 阈值为自由流行程时间1.5倍时的准时到达概率。

指标在每个时间步更新，采用 $O(1)$计算。

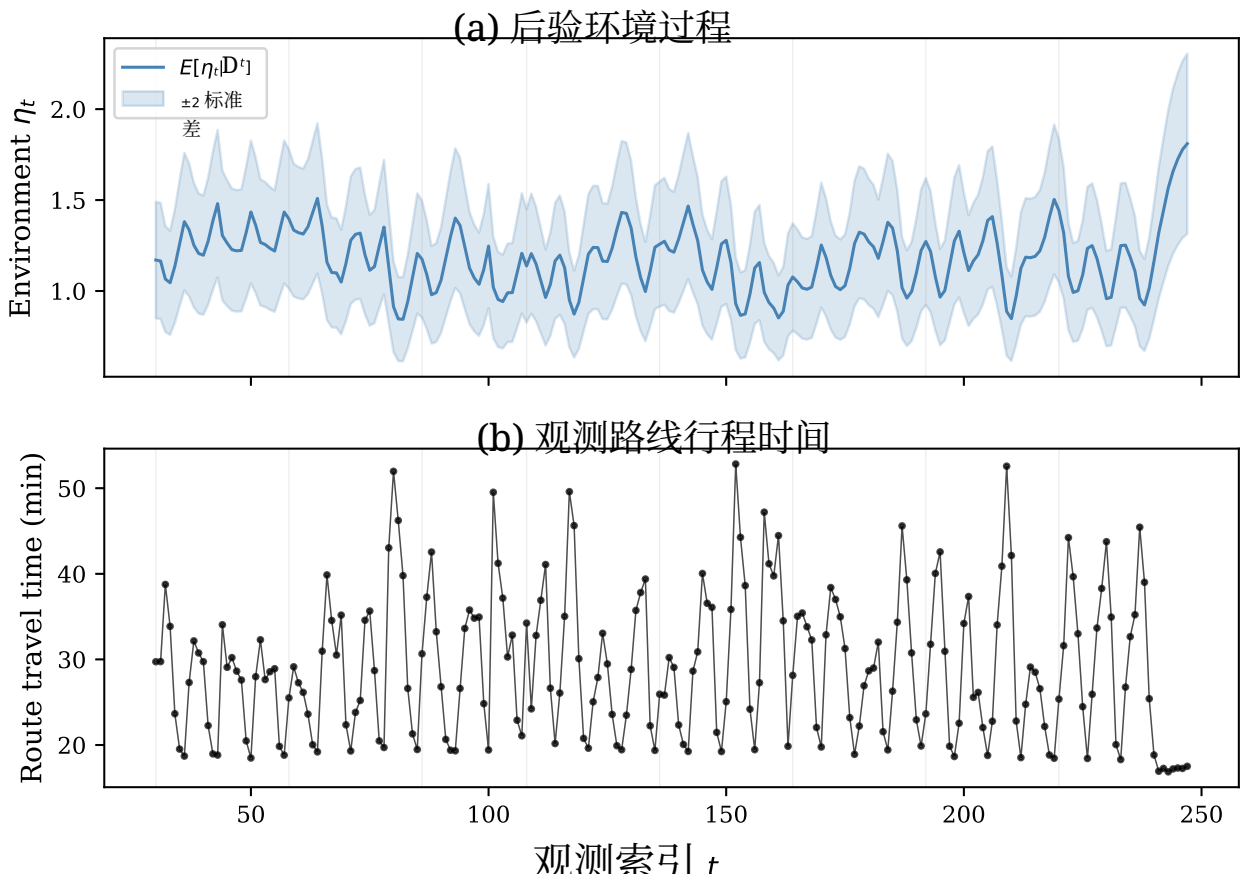

图 13：(a) 共同环境 η_t的后验均值和 ± 2 标准差带。数值越高表示路况越好（行程时间越短）。(b) 相应的观测路线行程时间。环境过程跟踪了观测到的拥堵动态。

图 13 展示了后验环境过程 η_t 以及观测路线行程时间。环境过程与拥堵程度呈反向变化：在自由流时段（行程时间较短）上升，在拥堵时段（行程时间较长）下降。这一潜在过程捕捉了同时影响所有 m 路段行程时间的共同路况，以简约的方式表征了该走廊的状态。

7.6 与替代方法的比较

我们框架的核心主张是，它能够解决相关的路段到路线聚合问题——这历来是实现路线可靠性闭式解的主要障碍——同时保持与假设独立性的方法相同的 $O(1)$计算成本。为此，我们将构建路线行程时间分布的四种替代方法进行比较：

1. **独立伽马卷积**（类似于 Lei 等，2014）：分别为每个路段拟合伽马分布，然后通过矩匹配得到路径分布，即

独立伽马分布之和。

2. **独立正态卷积**（类似于 Iida 1999）：计算各路段的均值和方差，并在独立假设下求和。

3. **具有伽马边缘分布的高斯 Copula**（类似于 Chen 等人 2017）：为各路段拟合伽马边缘分布，将其转换到正态空间，估计相关矩阵，并通过蒙特卡洛模拟（抽取 50,000 次）得到路线分布。

4. **静态伽马分布**（直接拟合）：对汇总后的路线行程时间数据拟合单一伽马分布。

与基于独立性的方法（1–2）的比较尤具参考意义，因为我们的模型处于相同的计算规模：两者都通过单个分布公式计算路线累积分布函数（独立性方法使用伽马分布，我们的方法使用 F-分布），无需模拟或数值积分。关键区别在于，我们的模型通过共同环境考虑了跨路段依赖，而基于独立性的方法则没有。Copula 方法（3）确实对依赖关系进行了建模，但需要进行计算成本高出数个数量级的蒙特卡洛模拟。此外，所有四种替代方法都是静态的：它们生成单一且不随时间变化的路线分布，而我们的模型则生成随时间变化的预测分布，并随着新数据到来而调整。

表 7：路线层面预测性能比较。KS p：概率积分变换均匀性的柯尔莫哥洛夫-斯米尔诺夫检验 p值（越高越好）。覆盖率：90% 预测区间内观测值所占比例（名义值为 0.90）。IW：90% 区间的平均宽度。

方法	动态	依赖性	KS p	90% 覆盖率	IW（分钟）
动态Gamma（我们的方法）	**Yes**	**Yes**	**0.467**	**0.954**	**31.5**
高斯 Copula +伽马分布	No	Yes	0.082	0.913	25.8
静态伽马分布（直接）	No	N/A	0.162	0.931	26.4
独立伽马卷积	No	No	0.000	0.344	9.4
独立正态卷积	No	No	0.000	0.367	10.1

表 7 展示了明显的排序。基于独立性的方法——伽马卷积和正态卷积——表现极差，对标称 90% 预测区间的覆盖率仅为 34–37%。其区间宽度只有 9–10 分钟，远远不够，因为这些方法忽略了跨路段依赖：在独立假设下对路段方差求和，会使路线方差被低估约 $\sum_{j,\,k}\mathrm{Cov}(y_j,\ y_k)/\sum_j \mathrm{Var}(y_j) \approx 2.7$ 倍。这种严重的覆盖不足表明，无论单个路段的建模效果有多好，都不能仅靠对相互独立的路段分布求和来解决路段到路线的聚合问题。

采用伽马边缘分布的高斯 Copula 能够正确捕捉依赖结构，覆盖率达到 91.3%，柯尔莫戈洛夫–斯米尔诺夫检验的 p 值为 0.082，较独立性假设有显著改善。然而，这也带来了高昂的计算成本：生成路线分布

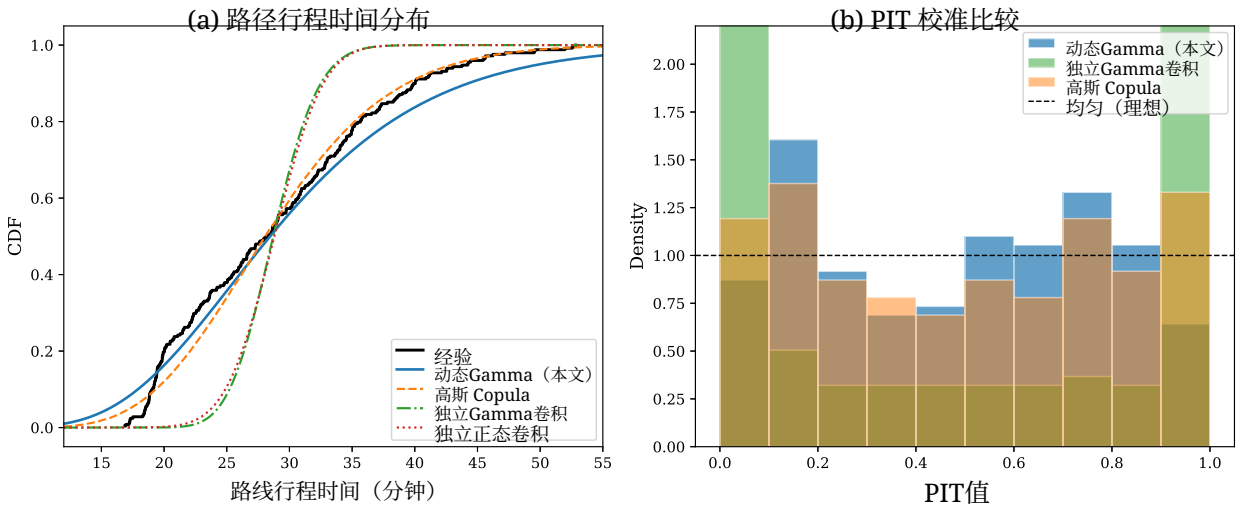

图 14：（a）路线行程时间累积分布函数：动态伽马模型（按时间段取平均）能够紧密跟踪经验累积分布函数，而基于独立性的方法所生成的分布过窄。（b）PIT 校准比较：只有动态伽马模型能够生成近似均匀的 PIT 值。

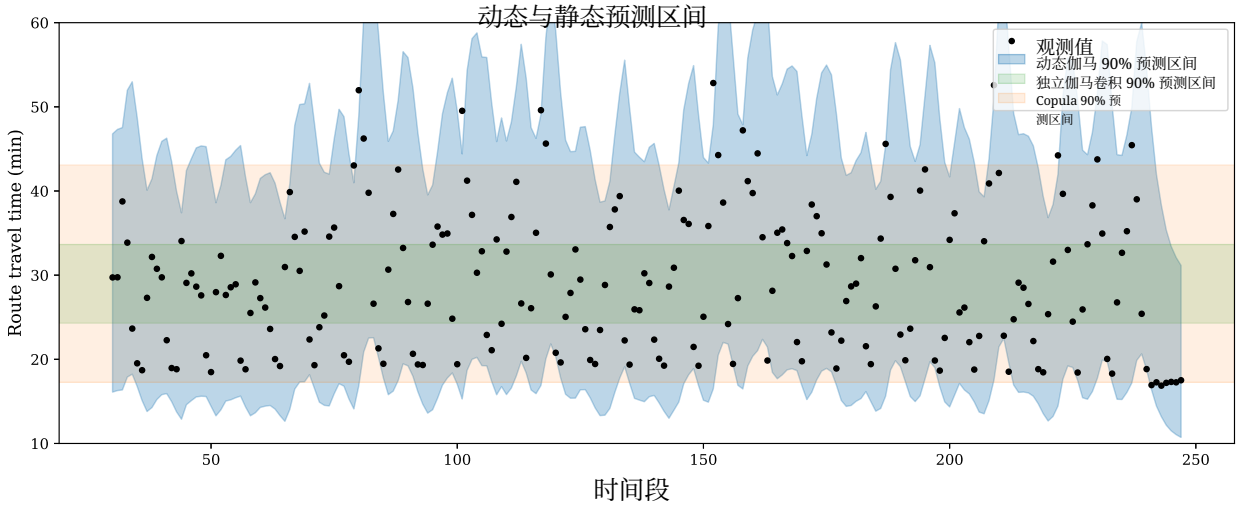

图 15：动态伽马模型的时变预测区间（蓝色带）与静态独立伽马卷积（绿色）和高斯 Copula（橙色）生成的恒定宽度区间的比较。动态模型会根据当前条件调整其不确定性，而静态方法要么覆盖不足（独立性方法），要么无法响应不断变化的交通状态（Copula）。

需要从 m 维高斯 Copula 中抽取 50,000 个样本，将每个样本通过路段特定的逆累积分布函数进行变换，并计算路线总和——在现代台式机上，这一过程耗时约 0.4 秒（根据我们的实现计时）。相比之下，我们的模型只需调用一次 F 分布（公式 14），即可计算路线累积分布函数，计算成本与基于独立性的方法相同。直接根据路线数据拟合的静态伽马分布表现尚可（覆盖率为 93.1%），但没有利用任何路段级信息，也无法组合用于备选路线。

我们的动态伽马模型实现了最佳的边际校准（名义柯尔莫戈洛夫–斯米尔诺夫检验 $p = 0.467$；参见第7节中的 PIT 自相关注意事项），覆盖率为 95.4%。关键在于，它在取得这一表现的同时仍能进行闭式计算：共同随机环境结构将跨路段依赖纳入条件独立框架，因此，相较于独立性假设，对依赖关系建模不会增加计算开销。时变区间在拥堵时段变宽，在自由流时段变窄（图 15）；无论如何处理依赖关系，任何静态方法都无法重现这一行为。

我们承认，这一比较混淆了模型的两项优势：(i) 纳入跨路段依赖；(ii) 生成随时间变化的预测。若要区分这两种效应，需要设置动态独立基线（例如，为每个路段分别拟合相互独立的动态伽马分布模型，并假设独立后进行卷积）。不过，我们指出，直接对路径数据拟合的静态伽马分布——它在汇总边际分布中隐式地捕捉了所有依赖关系——覆盖率仅为 93.1%，而我们的模型达到 95.4%。这表明，除依赖关系建模之外，时变成分也有显著贡献。我们还指出，评估设计并不对称：四种静态基线是在完整的 248 个观测值的样本上拟合并进行样本内评估，而动态伽马模型则仅使用截至 $t-1$ 的数据，生成真正的一步超前预测分布。事实上，这种不对称有利于静态基线；因此，它们表现较差就更能说明问题。另需说明，超参数（α，γ）通过经验贝叶斯（网格搜索）在用于评估的同一组 248 个观测值上选取；采用留出法或交叉验证评估会提供更严格的检验，不过，顺序进行的一步超前评估已经能够避免样本内过拟合。

8 讨论与结论

我们提出了一种具有共同随机环境的动态伽马模型，为路径级行程时间可靠性问题提供了有原则且计算高效的解决方案。其主要理论贡献在于证明，共同环境结构——这一结构在物理上反映了影响路线所有路段的共享交通状况——能够解决相关的路段到路线聚合问题，且不会带来任何计算开销：原本需要进行 m 维积分或蒙特卡洛模拟才能计算的路线可靠性指标，现在可以通过 F 分布以封闭形式计算，计算成本与仅假设独立性的方法相同，为 $O(1)$。第7节中的经验比较证实了这一点：

计算成本相同的基于独立性的方法，覆盖率仅为34–37%；而我们的模型通过正确考虑跨路段依赖，实现了95.4%的覆盖率。

主要结果如下。通过单一潜在过程 η_t，以简洁的方式捕捉跨路段依赖，无需显式进行联结函数建模或高维模拟。给定 η_t 后，各路段相互独立，因此路线分布问题可从 m 维降至一维。针对条件和的伽马矩匹配近似保留了共轭结构，从而使路径预测分布具有闭式伽马分布。最后，模型可按顺序以闭式形式更新，使其能够随着新传感器数据的到来实时跟踪可靠性。

对 I-55 的经验应用表明，该模型能够生成校准良好的路径级预测。超参数（α，γ）$=$(1.0, 0.70)通过经验贝叶斯方法选取，路段率参数 λ_j 通过马尔可夫链蒙特卡罗（采用前向滤波后向抽样的吉布斯采样）估计。模型生成的预测区间具有正确的覆盖率，并通过了边际PIT校准检验。 λ_j 的后验分布紧密集中在其最大似然估计值附近，贝叶斯预测分布与插件预测分布几乎完全相同，这证实数据提供了充分的信息，可直接在实时运营中使用共轭在线更新。值得注意的是，这些结果尤为突出，因为路径预测分布是从路段级多变量模型推导而来的——它并非另行拟合路线数据的独立模型，而是条件独立结构的推论。

我们的动态伽马模型与用于旅行时间可靠性预测的 GARCH 框架（Yang 和 Wu，2018）相关：二者都关注行程时间的时变波动性。在 GARCH 模型中，条件方差 σ_t^2 随过去的平方误差和过去的方差而变化，因此置信区间在拥堵时变宽，在自由流状态下变窄。在我们的模型中，折扣因子 γ 起到类似于 GARCH 持续性参数的作用：它控制过去信息衰减的速度，以及模型对近期冲击的响应程度。关键区别在于，我们的模型采用共轭贝叶斯框架，能够得到完整的预测分布——而不仅仅是围绕点预测的置信区间；此外，模型在路段层级运行，并通过共同环境结构实现向路径层级的组合。

未来值得研究几项扩展。首先，正如第 4 节所述，单一共同环境 η_t 意味着所有路段对之间都具有等相关性，而经验相关矩阵则表现出距离衰减。一个自然的扩展是引入多个环境过程——例如，每个子走廊一个，或每种拥堵状态一个——每个过程由一部分路段共享。这样将产生分块等相关结构，以更好地近似经验空间模式，但代价是需要增加潜在过程。同样，允许路段特定的形状参数 α_j（而非共同的 α）将能更好地捕捉路段层面的分布异质性，但会失去精确的共轭结构。其次，通过模型中的 $e^{\beta^T u_t}$ 项纳入观测协变量 u_t（天气、事故、特殊事件），可以改善非经常性状况下的预测。第三，Aktekin

等人（2020）提出的 Weibull扩展，其中 $y_{jt} \mid \eta_t \sim \text{Wei}(\lambda_j \eta_t,\ \phi_j)$，能够提高尾部行为的灵活性，但代价是失去具有封闭形式的路径预测分布。第四，允许路段参数 λ_j动态演化，可以捕捉路段特定的非平稳性，例如施工区域或反复出现的瓶颈形成。

第五，第7节（图5）中的静态混合分解揭示了路径行程时间明显的双峰性，分别对应自由流（$\mu = 19$ 分钟）和拥堵（$\mu = 31$ 分钟）分量。当前动态模型通过 η_t处理跨时间段的状态切换，并随着条件变化，将预测分布推向相应的峰值。不过，在单个时间段内，伽马分布假设是单峰的。一种自然的扩展是采用动态伽马模型混合，$p(y_t \mid \eta_t) = \sum_{k=1}^{K} \pi_k$ 伽马分布$(\alpha_k,\ \lambda_k \eta_t)$，其中每个分量对应一种交通状态。分量特定参数（$\alpha_k,\ \gamma_k$）至关重要：当所有分量共享相同的（$\alpha,\ \gamma,\ \lambda$）时，折扣因子会以 γ^t的速度消除初始化差异，导致各分量收敛到相同状态。共同环境 η_t 仍会驱动所有状态和路段的共享条件，而混合分布则捕捉时间段内的多峰性。路线分布的闭式表达将不再直接成立，因为条件和本身将是伽马分布的混合；要进行路径层面的推断，需要采用变分近似或基于模拟的近似。推断还需要通过潜在状态指标扩展模型，这些指标可通过吉布斯采样自然地处理。

最后，通过考虑共享公共路段且具有相关环境过程的多条路径，可以将该模型扩展到网络级可靠性，从而实现基于可靠性的网络交通管理。

参考文献

Aktekin, T., Polson, N. G. 和 Soyer, R.（2020）。一类多元非高斯时间序列模型。*Journal of Time Series Analysis*，41(5):691–721。Aktekin, T., Soyer, R. 和 Xu, F.（2013）。通过贝叶斯状态空间模型评估抵押贷款违约风险。《应用统计年鉴》，7(3):1450–1473。Carvalho, C., Johannes, M., Lopes, H. 和 Polson, N.（2010）。粒子学习与平滑。《统计科学》，25:88–106。Chen, C., Skabardonis, A. 和 Varaiya, P.（2003）。行程时间可靠性作为服务水平的衡量指标。《交通研究记录》，1855(1):74–79。Chen, M., Yu, G., Chen, P. 和 Wang, Y.（2017）。一种用于估计城市干道行程时间可靠性的基于联结函数的方法。《交通研究*C*辑：新兴技术》，82:1–23。Clark, S. 和 Watling, D.（2005）。随机需求下网络行程时间可靠性的建模。《交通研究*B*辑：方法论》，39(2):119–140。

Diebold, F. X., Gunther, T. A. 和 Tay, A. S.（1998）。密度预测评估及其在金融风险管理中的应用。*International Economic Review*，39(4):863–883。美国联邦公路管理局（2006）。行程时间可靠性：确保每次出行都准时抵达。技术报告，美国交通部。美国联邦公路管理局（2024）。用户定义出行路径上的旅行时间分布估计。技术报告 FHWA-HOP-20-066，美国交通部。Filipovska, M. 和 Mahmassani, H. S.（2021）。具有广义相关性的随机时变网络中可靠的轨迹自适应路径规划策略。《交通研究*C*辑：新兴技术》，133:103443。Fosgerau, M. 和 Fukuda, D.（2012）。旅行时间变异性的价值：城市道路旅行时间分布的特征。《交通研究*C*辑：新兴技术》，24:83–101。Fruhwirth-Schnatter, S.（1994）。数据增广与动态线性模型。*Journal of TimeSeries Analysis*，15(2):183–202。Gamerman, D. 和 Lopes, H. F.（2006）。用于贝叶斯推断的马尔可夫链蒙特卡罗随机模拟。Chapman & Hall/CRC。Hougaard, P.（2000）。多变量生存数据分析。Springer，纽约。Iida, Y.（1999）。道路网络可靠性分析的基本概念与未来方向。*Journalof Advanced Transportation*，33(2):125–134。Kamphuis, R., Mandjes, M. 和 Serra, P.（2025）。空间依赖网络中的道路交通估计与算法路径规划。*Scandinavian Journal of Statistics*，52(2):1058–1091。Kim, E. 和 Ye, H.（2022）。低频探测器数据下估计高速公路旅行时间的贝叶斯混合模型。*Applied Sciences*，12(13):6483。Lei, F., Wang, Y., Lu, G. 和 Sun, J.（2014）。具有不同服务水平的城市快速路行程时间可靠性模型。《交通研究*C*辑：新兴技术》，48:453–467。Lindley, D. V. 和 Singpurwalla, N. D.（1986）。共享共同环境的系统部件寿命的多变量分布。*Journal of Applied Probability*，23(2):418–431。Liu, H. X., Recker, W. 和 Chen, A.（2004）。利用实时环路检测数据揭示行程时间可靠性对动态路径选择的贡献。*Transportation Research Part A: Policy andPractice*，38(6):435–453。

Lomax, T., Schrank, D., Turner, S., and Margiotta, R. (2003). 选择出行可靠性度量指标。技术报告，德克萨斯交通研究所和 Cambridge Systematics。Mihaylova, L., Boel, R., and Hegyi, A. (2007). 粒子滤波框架下的高速公路交通估计。*Automatica*, 43(2):290–300。Moschopoulos, P. G. (1985). 独立伽马随机变量之和的分布。《统计数学研究所年刊》, 37(1):541–544。Nayak, T. K. (1987). 多元Lomax分布：其性质及在可靠性理论中的应用。*Journal of Applied Probability*, 24(1):170–177。Ng, M., Szeto, W., and Travis Waller, S. (2011). 利用概率不等式进行无分布假设的出行时间可靠性评估。《交通研究*B*辑：方法论》, 45(6):852–866。Ng, M. and Waller, S. T. (2010). 一种利用快速傅里叶变换刻画行程时间可靠性的计算高效方法。《交通研究*B*辑：方法论》, 44(10):1202–1219。Nie, Y. M., Wu, X., Dillenburg, J. F., and Nelson, P. C. (2012). 可靠路径引导：来自芝加哥的案例研究。*Transportation Research Part A: Policy and Practice*, 46(2):403–419。Polson, N. G. and Sokolov, V. O. (2015). 交通流的贝叶斯粒子跟踪。*IEEE Trans-actions on Intelligent Transportation Systems*, 16(5):2689–2699。Polson, N. G. and Sokolov, V. O. (2017). 用于短期交通流预测的深度学习。《交通研究*C*辑：新兴技术》, 79:1–17。Ramezani, M. and Geroliminis, N. (2012). 利用马尔可夫链估计城市主干道路线行程时间分布。《交通研究*B*辑：方法论》, 46(10):1576–1590。Saedi, R., Saeedmanesh, M., Zockaie, A., Saberi, M., Geroliminis, N., and Mahmassani, H. S. (2020). 利用网络划分估计网络行程时间可靠性。《交通研究*C*辑：新兴技术》, 112:46–61。Singpurwalla, N. D. (1995). 动态环境中的生存。《统计科学》, 10(1):86–103。Soyer, R., Aktekin, T., and Kim, B. (2016). 面向商业应用的计数时间序列贝叶斯建模。载于 Davis, R. A., Holan, S. H., Lund, R., and Ravishanker, N. 编，《离散值时间序列手册》，第275–296页。Chapman and Hall/CRC。Susilawati, S., Taylor, M. A., and Somenahalli, S. V. (2013). 城市道路行程时间变异性的分布。*Journal of Advanced Transportation*, 47(8):720–736。Taylor, M. A. (2013). 穿越时间：行程时间可靠性研究纪事。《*Transportmetrica B*：交通动力学》, 1(3):174–194。

Tebaldi, C., West, M., and Karr, A. F. (2002). 高速公路交通流的统计分析。《预测学杂志》, 21(1):39–68。Tilahun, N. Y. and Levinson, D. M. (2010). 时间的一刻：利用陈述偏好研究路径选择中的可靠性。《智能交通系统杂志》, 14(3):179–187。Tu, H., van Lint, H. W., and van Zuylen, H. J. (2007). 恶劣天气对高速公路走廊行程时间变异性的影响。载于交通研究委员会第*86*届年会，编号07-1642。van Lint, J., van Zuylen, H., and Tu, H. (2008). 高速公路上的旅行时间不可靠性：为何基于方差的度量只能揭示问题的一半。*Transportation Research Part A: Policy andPractice*, 42(1):258–277。West, M. and Harrison, J. (1997). 《贝叶斯预测与动态模型》。施普林格出版社，纽约，第2版。Westgate, B. S., Woodard, D. B., Matteson, D. S., and Henderson, S. G. (2016). 救护车的大型网络旅行时间分布估计。《欧洲运筹学杂志》, 252(1):322–333。Yang, S. and Wu, Y.-J. (2018). 使用ARIMA-GARCH模型对城市主干道行程时间进行时变置信区间预测。《韩国土木工程师学会土木工程杂志》, 22(10):4035–4044。Zang, Z., Xu, X., Qu, K., Chen, R., and Chen, A. (2022a). 交通网络中的行程时间可靠性：方法学进展综述。《交通研究*C*辑：新兴技术》, 143:103866。Zang, Z., Xu, X., Yang, C., and Chen, A. (2022b). 旅行时间分布估计综合综述。《*Transportmetrica B*：交通动力学》, 10(1):381–407。Zockaie, A., Nie, Y. M., and Mahmassani, H. S. (2014). 一种基于仿真的方法，用于在具有相关路段时间的随机网络中寻找最小行程时间预算路径。《交通研究记录》, 2467(1):140–148。