

基于数据驱动的青藏铁路冻土区行车可靠性分析模型^{*}

DATA-DRIVEN RELIABILITY ANALYSIS MODEL OF QINGHAI-TIBET RAILWAY IN PERMAFROST ZONE

王劲博^{**1} 唐家银^{***1,2} 刘新玲¹ 吴 怡¹

(1. 西南交通大学 数学学院, 成都 611756)

(2. 西南交通大学 综合交通大数据应用技术国家工程实验室, 成都 611756)

WANG JinBo¹ TANG JiaYin^{1,2} LIU XinLing¹ WU Yi¹

(1. School of Mathematics, Southwest Jiaotong University, Chengdu 611756, China)

(2. National Engineering Laboratory of Integrated Transportation Big Data Application Technology, Southwest Jiaotong University, Chengdu 611756, China)

摘要 为保障青藏铁路二期工程格尔木至拉萨段高原冻土区的行车安全,基于冻土区铁路路基原始时序高程值数据,建立不同路基类型下的乘积季节模型拟合路基不均匀沉降退化轨迹,同时置信化有效预测路基高程沉降值。通过起始高程值差分运算,进行路基沉降退化量统计分析,根据行车安全预设沉降阈值,借助失效时间外推方法完成可靠性评估。仿真模拟表明,模型能够有效地预测冻土区融沉退化情况且精度较高,同时研究成果可以为后勤保障部门的日常路基养护工作提供参考依据。

关键词 冻土融沉 乘积季节模型 退化数据 可靠性评估

中图分类号 U213.14 O213.2

Abstract To ensure the safety of train operation in the plateau permafrost region of the Golmud-Lhasa section of the Qinghai Tibet Railway Phase II project, based on the original time series elevation data of railway subgrade in permafrost region, the product season model under various subgrade types is established to fit the uneven settlement degradation trajectory of subgrade and effectively predicted elevation values with confidence intervals. The differential operation was used to analyze the subgrade settlement degradation values. The failure time extrapolation method was used to complete the reliability evaluation, and the settlement threshold was set based on driving safety. The simulation results show that it can accurately forecast permafrost thawing degradation, and the research findings can be used by the logistics support department to conduct daily subgrade maintenance.

Key words Permafrost thawing; SARIMA model; Degradation data; Reliability assessment

Corresponding author: TANG JiaYin, E-mail: tangjiayin@home.swjtu.edu.cn, Fax: +86-28-66367624

The project supported by the Humanities and Social Science Research Planning Fund Project of the Ministry of Education (No. 20XJAZH009), the Sichuan Province 2021-2023 Higher Education Talent Training Quality and Teaching Reform Project (No. JG2021-394), the Southwest Jiaotong University Undergraduate Education and Teaching Research and Reform Project (No. 20220320), and the 2021 Annual Project of Basic Research and Cultivation of Central Universities (No. 2682021ZTPY018).

Manuscript received 20220822, in revised form 20230104.

0 引言

青藏铁路作为中国新世纪四大工程之一,成功打通了内陆与西藏腹地之间的联系,是世界上高原铁路

的海拔之最,从 20 世纪 50 年代规划修建一期工程,前后经历半个世纪,直至 21 世纪初才完成全线通车。青藏铁路格尔木至拉萨段二期工程覆盖大面积高原冻土区,恶劣的自然环境和复杂的工程地质条件必然导致

^{*} 20220822 收到初稿, 20230104 收到修改稿。教育部人文社会科学研究规划基金项目 (20XJAZH009), 四川省 2021-2023 年高等教育人才培养质量和教学改革项目 (JG2021-394), 西南交通大学教育教学研究与改革项目重点项目 (20220320), 中央高校基础研究培育专题 2021 年度项目 (2682021ZTPY018) 资助。

^{**} 王劲博, 男, 1999 年生, 四川成都人, 汉族, 西南交通大学在读硕士研究生, 主要研究方向为应用统计、可靠性统计。

^{***} 唐家银 (通信作者), 男, 1976 年生, 安徽无为, 汉族, 西南交通大学副教授, 博士研究生导师, 主要研究方向为应用统计、可靠性理论与工程。

列车运营过程中呈现大量路基病害及潜在风险,多年冻土区的含冰量极高,暖季路基融沉及寒季路基冻胀势必导致铁路路基不均匀沉降,且发生频率和变形程度远高于其他线路,导致青藏铁路养护难度很大。

迄今为止,不少学者对青藏铁路高原冻土区路基变形的规律和路基力学模型进行了大量的研究。基于青藏铁路多年冻土区路基地温与变形现场监测资料,高宝林等^[1]探究了青藏铁路路基下融化夹层特征及其对路基沉降变形的影响;汪双杰等^[2]基于青藏高原冻土区公路路基病害进行研究,制定病害预防的安全公路行车工程相应对策;在青藏铁路冻土区路桥过渡段,王进昌等^[3]根据工程地质条件理论研究分析路基沉降特点,详细解释路基沉降成因;洪兆阳等^[4]通过利用创新的时序 PS-InSAR 监测方法,得出有关青藏高原冻土区地表形变过程的深刻结论;余雨薇等^[5]考虑环境温度与行车振动应力等解释因子,从统计角度开展青藏铁路路基沉降可靠性分析,并且通过实例表明可靠性评估基本满足工程需要。多数相关文献集中于路段路基的地质、土壤、水量和力学模型研究,而鲜见有基于现场铁路路段的监测数据,以数据驱动开展研究。由于路基变形工况呈现明显的趋势性-季节性综合效应特征,故在冻土区铁路路基沉降研究中,季节性效应分析是不可或缺的关键一环,运用乘积季节模型(Seasonal Autoregressive Integrated Moving Average, SARIMA)比较合理,目前已有关于基于 SARIMA 模型预测趋势的文献^[6-7]。

国内外学者对于冻土变形的大多数研究仅考虑数据的趋势性特征,缺少对季节性综合特征的研究。因此,本文一方面考虑路基类型产生的变形差异效应;另一方面,对路基变形的趋势性-季节性综合效应特征进行综合分析。基于铁路路基失效时间分布的退化数据统计分析,建立冻土区路基不均匀沉降预测模型,计算输出可靠度评价体系,服务于冻土路基养护技术管理办法措施,并进行青藏铁路路基预警,有助于冻土区铁路运营的养护保障工作。

1 冻土区路基不均匀沉降预测模型

1.1 乘积季节模型理论基础

时间序列分析基本思想是根据系统有限长度的观察数据,建立能较精准地反应序列中所包含动态依存关系的模型,以此对系统的未来进行预测。因此时间序列分析不仅可以揭示某一现象的发展规律,同时能从动态上刻画若干现象之间的内部变换规律,还可以预测和控制未来的现象,修正或重新设计系统。

季节效应是指某序列由于受自然气候、生产条件、生活习惯等因素的影响,在一定的时间内随季节的变

化而呈现出周期性变化的规律。基于时间序列中长期趋势效应、季节效应和随机波动之间的干涉作用,简单乘积季节模型未考虑不同周期同一周期点之间的相关性,并不能有效分离彼此,而且模型中的差分和季节差分并不能完全消除季节性,因此考虑使用乘积季节 SARIMA 模型,即模型采用乘法方式对季节差分与差分进行组合。

SARIMA 模型中存在大量参数:时间序列与白噪声序列分别记为 $\{X_t\}$ 与 $\{\varepsilon_t\}$, B 代表后移算子, p, q 分别代表消除同一周期不同周期点之间的相关性的自回归阶数及移动平均阶数, s 代表周期长度, P, Q 分别代表消除不同周期同一周期点之间的相关性的自回归阶数及移动平均阶数, d, D 分别代表差分阶数及季节差分阶数。基于季节差分和差分结果,通过上述参数的多项式组合, X_t 与 $X_{t-1}, X_{t-2}, \dots, X_{t-p}, X_{t-s}, X_{t-2s}, \dots, X_{t-ps}$ 及 $\varepsilon_t, \varepsilon_{t-1}, \varepsilon_{t-2}, \dots, \varepsilon_{t-q}, \varepsilon_{t-s}, \varepsilon_{t-2s}, \dots, \varepsilon_{t-Qs}$ 相关, SARIMA 模型对于随机序列存在如下结构^[8]:

$$\varphi(B)U(B^s)(1-B)^d(1-B^s)^D X_t = \theta(B)V(B^s)\varepsilon_t \quad (1)$$

SARIMA 模型中白噪声序列 $\{\varepsilon_t\}$ 满足 Gauss-Markov 条件,即零均值、同方差和不相关,并且 ε_t 仅对 X_t 有影响,与 $X_k, k < t$ 无关,其表达式为

$$\begin{cases} E(\varepsilon_t) = 0 \\ \text{Cov}(\varepsilon_t, \varepsilon_k) = \begin{cases} \sigma_\varepsilon^2, t = k \\ 0, t \neq k \end{cases} \\ E(X_k \varepsilon_t) = 0, \forall k < t \end{cases} \quad (2)$$

SARIMA 模型中平稳可逆项和季节项的自回归系数多项式 $\varphi(B), U(B^s)$ 及移动平均系数多项式 $\theta(B), V(B^s)$ 具体表达式为

$$\begin{cases} \varphi(B) = 1 - \varphi_1 B - \varphi_2 B^2 - \dots - \varphi_p B^p \\ \theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q \\ U(B^s) = 1 - u_1 B^s - u_2 B^{2s} - \dots - u_P B^{Ps} \\ V(B^s) = 1 - v_1 B^s - v_2 B^{2s} - \dots - v_Q B^{Qs} \end{cases} \quad (3)$$

SARIMA 模型中后移算子 B 作用一次 X_t 或 ε_t 后使其转换为 X_{t-1} 或 ε_{t-1} , 目的在于将时间序列和白噪声序列当期值转换为过去时期值,其表达式为

$$\begin{cases} \varphi(B)X_t = X_t - \varphi_1 X_{t-1} - \dots - \varphi_p X_{t-p} \\ \theta(B)\varepsilon_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q} \\ U(B^s)X_t = X_t - u_1 X_{t-s} - \dots - u_P X_{t-Ps} \\ V(B^s)\varepsilon_t = \varepsilon_t - v_1 \varepsilon_{t-s} - \dots - v_Q \varepsilon_{t-Qs} \end{cases} \quad (4)$$

SARIMA 模型中弱趋势项的差分多项式 $(1-B)^d$ 及季节项的季节差分多项式 $(1-B^s)^D$ 具体表达式以 $d=1, D=1$ 和 $d=2, D=1$ 为例:

$$\begin{cases} (1-B)^1(1-B^s)^1X_t = \\ (X_t - X_{t-1}) - (X_{t-s} - X_{t-1-s}) \\ (1-B)^2(1-B^s)^1X_t = \\ (X_t - 2X_{t-1} + X_{t-2}) - (X_{t-s} - 2X_{t-1-s} + X_{t-2-s}) \\ \dots \end{cases} \quad (5)$$

SARIMA 模型实质上是自回归整合移动平均 ARIMA(p, d, q) 模型的一般化推广, 简记为 ARIMA(p, d, q) $\times$ (P, D, Q, s), SARIMA 模型采用 Box-Jenkins 方法建立已经相当成熟。在实际问题中, 季节性时间序列所含有的成分不同且记忆性长度各异, 因而模型形式也是多种多样的。对工程问题中的时序数据处理, 基于月度及季度的概念, 季节周期 $s = 4$ 或 $s = 12$ 是普遍现象。

1.2 基于乘积季节模型预测冻土区路基沉降值

乘积季节模型最小均方误差预测的概念: 假设当前时刻为 t , 则使用当前及过去观测值 $X_t, X_{t-1}, X_{t-2}, \dots, X_{t-p}, X_{t-s}, X_{t-2s}, \dots, X_{t-ps}$ 对未来 k 步长时刻 $t+k$ 处观测值 X_{t+k} 作出预测, 预测值记为 $\hat{X}_t(k)$ 是过去时间序列值的函数, 预测误差记为 $\hat{e}_t(k)$, 最小均方误差表示为

$$\min E[\hat{e}_t(k)]^2 = \min E[X_{t+k} - \hat{X}_t(k)]^2 \quad (6)$$

SARIMA 模型最小均方误差预测的推导过程: 将原始时间序列 $\{X_t\}$ 通过差分多项式 $(1-B)^d$ 与季节差分多项式 $(1-B^s)^D$ 达到平稳性的同时消除部分季节性, 将处理后的新时间序列记为 $\{X'_t\}$, 引入格林函数 $G(B)$ 的概念^[9], SARIMA 模型传递形式简化为

$$X'_t = \varphi^{-1}(B)U^{-1}(B^s)\theta(B)V(B^s)\varepsilon_t = G(B)\varepsilon_t \quad (7)$$

格林函数 $G(B)$ 意义在于时间序列通过线性变换表示成既往白噪声的加权求和形式, 将下标转换为 $t+k$ 后, SARIMA 模型传递形式具体为

$$X'_{t+k} = G(B)\varepsilon_{t+k} = G_0\varepsilon_{t+k} + \dots + G_{k-1}\varepsilon_{t+1} + \sum_{j=0}^{\infty} G_{k+j}\varepsilon_{t-j} \quad (8)$$

由于预测只能建立在到 t 时刻为止的可用信息的基础上, 引入权系数 D_j , 其在最小均方误差的意义下确定为

$$\hat{X}'_t(k) = D_0\varepsilon_t + D_1\varepsilon_{t-1} + \dots = \sum_{j=0}^{\infty} D_j\varepsilon_{t-j} \quad (9)$$

规定 t 为时间原点, 向前 k 步的预测误差表示为

$$\begin{aligned} \hat{e}_t(k) &= X'_{t+k} - \hat{X}'_t(k) = G_0\varepsilon_{t+k} + \dots + G_{k-1}\varepsilon_{t+1} + \\ &\sum_{j=0}^{\infty} (G_{k+j} - D_j)\varepsilon_{t-j} \end{aligned} \quad (10)$$

因为预测误差需要满足均值为零, 所以最小均方误差等价于误差的最小方差, 白噪声序列具有同方差 σ_ε^2 。

$$\begin{aligned} E[\hat{e}_t(k)]^2 &= \text{Var}[\hat{e}_t(k)] = \\ &\sigma_\varepsilon^2 \left[\sum_{j=0}^{k-1} G_j^2 + \sum_{j=0}^{\infty} (G_{k+j} - D_j)^2 \right] \end{aligned} \quad (11)$$

当 $G_{k+j} = D_j$ 时误差的方差达到最小即最小均方误差, 同时实际工程问题中数据维度是有限的, 基于格林函数的指数衰减性或者截尾性保证近似效果, X'_{t+k} 的最小均方误差预测为

$$\begin{aligned} \hat{X}'_t(k) &= G_k\varepsilon_t + G_{k+1}\varepsilon_{t-1} + G_{k+2}\varepsilon_{t-2} + \dots = \\ &\sum_{j=0}^{\infty} G_{k+j}\varepsilon_{t-j} = \sum_{j=0}^T G_{k+j}\varepsilon_{t-j} \end{aligned} \quad (12)$$

$\hat{X}'_t(k)$ 最后完成差分多项式 $(1-B)^d$ 与季节差分多项式 $(1-B^s)^D$ 的组合即可得到 X_{t+k} 的最小均方误差预测 $\hat{X}_t(k)$ 。

通过上述推导, 最小均方误差预测的方差和预测步长 k 有关, 而与预测的时间原点 t 无关; 同时预测步长 k 越大, 预测误差的方差也越大, 即预测的准确性越差。

基于部分已知信息获得的预测值 $\hat{X}_t(k)$, 对比剩余真实值 X_{t+k} 得到统计距离, 检验退化模型的预测精度。选取相同量纲的均方根误差 (Root Mean Square Error, RMSE) 和平均绝对误差 (Mean Absolute Error, MAE) 作为预测评价指标, 当两者均处于较低水平时, 预测效果值得信赖, 其表达式为

$$\begin{cases} X_{\text{RMSE}} = \sqrt{\frac{1}{n} \sum_{k=1}^n (\hat{X}_t(k) - X_{t+k})^2} \\ X_{\text{MAE}} = \frac{1}{n} \sum_{k=1}^n |\hat{X}_t(k) - X_{t+k}| \end{cases} \quad (13)$$

1.3 基于冻土区路基预设沉降阈值的安全行车周期预测

目前工程中可靠性数据分析方法都是基于失效时间数据的, 但是在实际中还存在一类数据: 与可靠性直接相关的某个性能指标随时间推移呈现具有一定趋势的变化规律, 并且该性能指标本身是可以测量的, 这种现象定义为退化过程, 具有这种现象的性能指标称为性能退化量^[10]。

假设在 t 时刻的退化量为 $y(t)$, 退化失效可定义为当退化量随时间变化达到预定失效标准 D_t 时立即发生失效, 则失效时间 T_0 对应表示为^[11]: $T_0 = \inf\{t: y_t = D_t; t \geq 0\}$ 。

高原冻土区出现融沉现象较为频繁, 铁路路基大多产生不均匀沉降, 规定 t 为时间原点, 在将来 $t+k$ 时刻的沉降退化量预测值 $\hat{y}_t(k)$, 规定 SARIMA 模型中的时序初始值与时序预测值分别为 X_0 与 $\hat{X}_t(k)$, 那么

$$\hat{y}_i(k) = X_0 - \hat{X}_i(k)。$$

沉降退化量 $\hat{y}_i(k)$ 的预测结果仅存在一个具体的数值,虽然便于计算和使用,但是不能反映自身精度。当确定显著性水平 α 后,引入置信区间的概念^[12],构造置信下限及置信上限两个统计量,其中 u_p 代表标准正态分布中的 p 分位数。平均而言,基于大量的区间估计观测值中,至少有 $100(1 - \alpha)\%$ 包含预测结果 $\hat{y}_i(k)$,其表达式为

$$\begin{cases} \hat{y}_{i_{\text{lower}}}(k) = \hat{y}_i(k) - u_{1-\frac{\alpha}{2}} \sqrt{\text{Var}[\hat{e}_i(k)]} \\ \hat{y}_{i_{\text{upper}}}(k) = \hat{y}_i(k) + u_{1-\frac{\alpha}{2}} \sqrt{\text{Var}[\hat{e}_i(k)]} \end{cases} \quad (14)$$

基于沉降退化量 $\hat{y}_i(k)$ 与失效标准 D_f 的定义,安全行车周期即可推导。同时基于沉降退化量 $\hat{y}_i(k)$ 的置信上下限,安全行车周期的置信区间估计也可完成,进一步保障高原冻土区列车行驶安全。

2 既定变形阈值下安全行车可靠性评估

2.1 冻土区路段路基失效时间分布拟合

对于具有退化特征的安全性评估对象,其具有随机波动的退化轨迹达到安全阈值极限所需要的时间也具有某种随机性。因此利用 Weibull 分布来描述冻土区铁路路基不均匀沉降过程的失效时间,进而对行车安全的可靠性水平进行定量评估,其概率密度函数与累积分布函数定义为^[13]

$$\begin{cases} f(t) = \frac{m}{\eta} \left(\frac{t - \gamma}{\eta} \right)^{m-1} e^{-((t-\gamma)/\eta)^m} \\ F(t) = 1 - e^{-((t-\gamma)/\eta)^m} \end{cases} \quad (15)$$

Weibull 型分布属于连续型分布,其分布特征由形状参数 m 、尺度参数 η 和位置参数 γ 决定。形状参数 m 规定概率密度函数的趋势走向;尺度参数 η 作用于坐标尺度的伸缩,在列车运营期间直观反映出铁路路基负载水平;位置参数 γ 在数学意义上代表函数的横坐标平移变换,实际工程中标志路基沉降导致的失效时间起点。

基于 Weibull 分布在工程应用上的广泛性和一般性,运用回归识别拟合威布尔分布参数,不同于极大似然估计求偏导时的复杂计算,运用回归分析中最小二乘估计方法(表 1)^{[14]19-22}。

操作步骤中需注意:首先,经验分布的函数图像为非降右连续的阶梯函数,当样本容量 $n \geq 30$ 时,满足样本经验逼近总体理论分布的条件^{[15]82-83},基于回归拟合优度指标样本决定系数^{[14]35-36} r^2 评估回归效果,当 $r^2 \geq 95\%$ 时拟合程度较高;其次,由于 $\ln(1 - F(t))$ 的定义域非负,适当修正经验分布 $F_n(t_{(i)}) = (i - 0.3)/(n + 0.4)$,避免 $F(t) = 1$ 时函数无界,同时修正因子数值较低,不会显著改变分布特征;最后,解释变量 X 中仍然包含未知参数 γ ,因为位移参数实际意义即产路基失效起点,所以可以选择样本路基失效时间中的最小值 $t_{(1)}$ 。

2.2 冻土区行车安全可靠评估

冻土区安全行车可靠度指铁路路基在规定时间内和频发融沉现象的条件下,完成列车平稳行驶的能力,通常记为 $R(t)$;安全行车失效率表示运营到 t 时刻列车始终保持安全行驶,基于 t 时刻后,单位时间内

表 1 参数的最小二乘估计

Tab. 1 Least squares estimation of the parameter

步骤 Step	具体操作 Specific operation
1	基于 SARIMA 模型预测的路基失效时间样本 $T: \{\hat{t}_1, \hat{t}_2, \dots, \hat{t}_n\}$ 获得次序统计量,从小到大排列得到次序序列 $t_{(1)} \leq t_{(2)} \leq \dots \leq t_{(n)}$ Sequential statistics were obtained based on subgrade failure time sample $T: \{\hat{t}_1, \hat{t}_2, \dots, \hat{t}_n\}$ predicted by SARIMA model, and arrange from smallest to largest to get an ordered sequence $t_{(1)} \leq t_{(2)} \leq \dots \leq t_{(n)}$
2	按照次序统计量构造经验分布函数: $F_n(t_{(i)}) = i/n$, Glivenko-Cantelli ^{[15]80-81} 定理表明样本量充足时,经验分布函数 $F_n(x)$ 很好地近似于总体分布函数 $F(x)$ Construct the empirical distribution function according to the order statistic: $F_n(t_{(i)}) = i/n$, the Glivenko-Cantelli theorem shows that when the sample size is sufficient, the empirical distribution function is well approximated to the population distribution function
3	$F(x)$ 做理论推导转换成线性回归模型: $\ln\{1 - F(t)\} = -m \ln \eta + m \ln(t - \gamma)$ Theory derives the transformation of $F(x)$ into a linear regression model: $\ln\{1 - F(t)\} = -m \ln \eta + m \ln(t - \gamma)$
4	$\ln\{1 - F(t)\}$ 等价于被解释变量 Y , $\ln(t - \gamma)$ 看作解释变量 X , $-m \ln \eta$ 和 m 类比成线性回归模型中的参数 β_0 和 β_1 , 那么上式可以化简为 $Y = \beta_0 + \beta_1 X + \varepsilon$ $\ln\{1 - F(t)\}$, $\ln(t - \gamma)$, $-m \ln \eta$ and m are equivalent to the explanatory variable Y , X , β_0 and β_1 , then the above equation can be reduced to: $Y = \beta_0 + \beta_1 X + \varepsilon$
5	Weibull 分布的参数估计转换为运用最小二乘法估计线性回归模型参数,其中 $m = \beta_1$, $\eta = \exp(-\beta_0/\beta_1)$, $\gamma = t_{(1)}$ 可操作性显著增加 The parameter estimation of the Weibull distribution is converted to the estimation of linear regression model parameters using the least squares method, where the operability of $m = \beta_1$, $\eta = \exp(-\beta_0/\beta_1)$ and $\gamma = t_{(1)}$ is significantly increased

路基发生失效的频率,通常记为 $\lambda(t)$; 平均安全行车周期 $E(T)$ 直观表现路基失效时间 T 均值,常被铁路养护单位使用;安全行车的可靠信息 t_R 代表保证铁路路基正常工作的概率在特定水平 R 以上,列车运营可

以正常工作的时间。上述重要的可靠性指标在 Weibull 型分布条件下,存在具体的函数表达式^[16-17],详情见表 2。

表 2 Weibull 型分布可靠性指标
Tab. 2 Weibull type distribution reliability index

可靠度函数 Reliability function	失效率函数 Failure rate function	平均安全行车周期 Mean driving safety cycle	安全行车可靠信息 Reliable information for driving safety
$R(t) = \int_t^\infty f(\tau) d\tau = e^{-((t-\gamma)/\eta)^m}$	$\lambda(t) = \frac{m}{\eta^m} (t - \gamma)^{m-1}$	$E(T) = \gamma + \eta \Gamma(1 + 1/m)$	$t_R = \gamma + \eta (-\ln R)^{1/m}$

需要注意可靠度函数 $R(t)$ 是通过概率密度函数 $f(\tau)$ 无穷积分得到;失效率函数 $\lambda(t)$ 基于 m 的数值,将路基的不同失效类型区分为递增型 ($m > 1$)、恒定型 ($m = 1$) 及递减型 ($m < 1$); 平均安全行车周期 $E(T)$ 中的 $\Gamma(x)$ 代表 Gamma 函数,具体形式^[18]为 $\Gamma(x) = \int_0^{+\infty} t^{x-1} e^{-t} dt$; 安全行车可靠信息 t_R 随特定水平 R 的提升而下降。

3 K1425+000~K1569+000 路段路基不均匀沉降研究分析

3.1 高原冻土区各路基类型预处理

东起青海省格尔木市,西至西藏自治区拉萨市的青藏铁路二期工程全长 1 142 km,设计最高行车速度 100 km/h。路基相对格尔木海拔高程数据获取源于 K1425+000~K1569+000 路段非等距安插的 295 观测点在 2010 年 1 月至 2019 年 10 月内的记录情况。青

藏铁路二期工程位于高原冻土区的铁路路基类型包括十种,其中一般路基、通风片石和一般路基、片石护道类型各仅存在一处观测点,其样本容量过低不具有统计价值,其余路基类型具体数量记录于表 3。

事实上不同的路基类型对铁路路基维护效果存在显著差距,例如:热棒是一种由碳素无缝钢管制成的高效热导装置,具有独特的单向传热性能,热量只能从地面下端向地面上端传输,反向不能传热,能够有效防治青藏铁路冻土路基融沉。

3.2 各路基类型退化轨迹预测情况

对各路基类型样本中进行简单随机抽样单观测点,运用 Sarima 模型建立非线性退化模型,服务于预测未来铁路路基高程值,差分运算后即可获得路基沉降退化量。按照 Box-Jenkins 建模方法步骤直至完成诊断检验后,各路基类型的 Sarima 模型参数如表 3 所示。

表 3 各路基类型 Sarima 模型具体参数
Tab. 3 Specific parameters of Sarima model for each subgrade type

路基类型 Subgrade type	样本容量 Sample size	标记 Tag	抽样点 Sampling point	模型阶数 Model order
热棒、碎石护坡、片石护道 Thermosyphon, rubble slope and gravel pavement	24	A 类 Class A	K1425+000	ARIMA(0,1,2) × (0,0,0,0)
碎石护坡、片石护道 Rubble slope and gravel pavement	43	B 类 Class B	K1425+100	ARIMA(0,2,2) × (1,0,0,12)
一般路基、骨架护坡 General subgrade and brick skeleton slope	97	C 类 Class C	K1427+200	ARIMA(0,1,1) × (0,0,1,12)
一般路基 General subgrade	52	D 类 Class D	K1447+400	ARIMA(0,1,0) × (0,0,0,0)
碎石护坡 Rubble slope	18	E 类 Class E	K1458+950	ARIMA(1,1,1) × (0,0,0,0)
骨架护坡、片石护道 Brick skeleton slope and gravel pavement	15	F 类 Class F	K1475+750	ARIMA(1,0,0) × (0,0,1,12)
骨架护坡、一般路基 Brick skeleton slope and general subgrade	19	G 类 Class G	K1496+500	ARIMA(2,1,1) × (0,0,0,0)
骨架护坡、一般路基、碎石护坡、护道 Brick skeleton slope, general subgrade, rubble slope and pavement	25	H 类 Class H	K1568+000	ARIMA(2,1,2) × (0,0,0,0)

基于定阶后的 Sarima 模型,预测未来路基高程

值,同时考虑路基不均匀沉降过程中存在测量误差、极

端天气出现,以及列车载荷差异带来的随机效应影响。若使用点估计高程预测值,直接将抽样指标代替全体指标,预测结果将过于片面;而区间估计通过构造枢轴量的方法寻找优良的置信区间,概念上既回答高程预

测值相较于真值的距离,又能提供估计精度同时涵盖噪声项 $\{\varepsilon_t\}$, 预测结果的将可信程度大幅提升。规定显著性水平 $\alpha = 0.02$, 给出对应的预测置信区间,在图 1 中使用阴影覆盖部分表示。

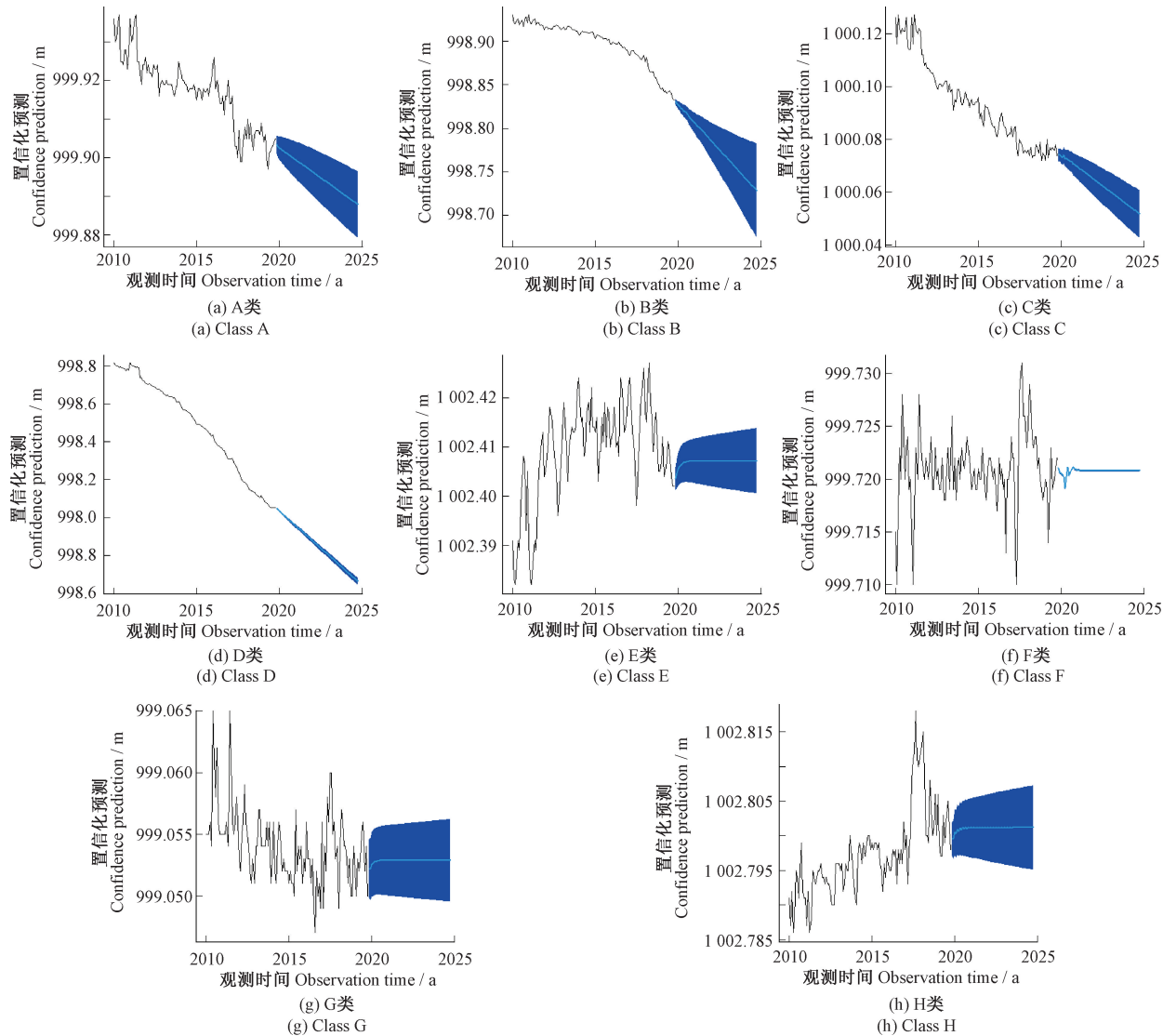


图 1 各路基类型高程区间估计预测情况

Fig. 1 Estimation and prediction of the elevation interval of each subgrade type

作为路基不均匀沉降预测模型,各路基类型的 SARIMA 模型的精度亟待验证。由于单退化轨迹样本容量相对充裕,视 2010 年 1 月至 2018 年 12 月路基高程值数据为训练集实现退化模型的建立,将 2019 年 1 月至 2019 年 10 月作为测试集。这样,前者产生的预测结果能够与后者的真实情况进行比较,使用预测评价指标 RMSE 及 MAE,衡量各路基类型的 SARIMA 模型预测性能,详情见表 4。

预测评价指标数值在可接受的范围内,说明预测值相当接近真值,各路基类型的 SARIMA 模型预测精度较高,后续的安全行车周期外推及行车可靠性评估得以保证。

表 4 不均匀沉降预测模型性能评价指标
Tab. 4 Performance evaluation index of uneven settlement prediction model

路基类型 Subgrade type	均方根误差 Root mean square error X_{RMSE} /mm	平均绝对误差 Mean absolute error X_{MAE} /mm
A 类 Class A	3. 81	2. 95
B 类 Class B	10. 87	9. 45
C 类 Class C	4. 96	4. 23
D 类 Class D	9. 03	8. 31
E 类 Class E	5. 74	4. 72
F 类 Class F	3. 06	2. 28
G 类 Class G	1. 92	1. 64
H 类 Class H	3. 77	3. 39

3.3 基于路基不均匀沉降外推安全行车周期

格尔木至拉萨 K1449+350~K1450+300 路段,路基类型:一般路基、骨架护坡,纵观整个二期工程在 2010 年 1 月至 2019 年 10 月间融沉现象较为严重,保证安全行车的路基沉降情况作为性能退化量显著上升,逐渐逼近失效阈值。通过青藏铁工桥[2011]12 号多年冻土路基养护技术管理办法,路基坡面下沉或外鼓,造成路肩错台,且错台高度超过 0.4 m 应立即封锁线路,即沉降量的失效标准 $D_f = 0.4$ m。同时,紧急整修后放行列车不超过 45 km/h。

上述路段包含 33 处观测点,样本容量达到可靠性数据分析基础,将 2010 年 1 月规定为起始时间,当月路基高程值标记为 X_0 。依次按照退化轨迹建模步骤,预测得到沉降幅值的点估计、区间估计置信下限及置信上限,继而标记首次达到失效阈值 D_f 的时间样本分别为 $\{\hat{t}_i\}$ 、 $\{\hat{t}_{iL}\}$ 和 $\{\hat{t}_{iU}\}$,其中 $i = 1, 2, \dots, 33$,失效时

间等价于安全行车周期。

3.4 K1449+350~K1450+300 路段行车安全可靠推断

基于样本铁路路基安全行车周期样本,相应的升序排列结果分别记为 $\{\hat{t}_{(i)}\}$ 、 $\{\hat{t}_{(iL)}\}$ 和 $\{\hat{t}_{(iU)}\}$ 。对应构造修正经验分布函数,继而按照最小二乘估计原则,在回归拟合优度指标样本决定系数符合既定标准上,完成 Weibull 分布的参数估计,详情见表 5。

基于参数估计的结果,可以初步定性评估铁路工程冻土区 K1449+350~K1450+300 路段路基不均匀沉降的可靠性:可靠度函数 $R(t)$ 与失效率函数 $\lambda(t)$ 均随时间变化,其中可靠度函数是单调性满足递减,失效率函数单调性服从递增。结论表明,路基均匀沉降导致可靠度显著下降,单位时间内失效频率显著上升。根据表 2 提供的可靠性指标定量评估可靠性,详情见表 6 与图 2。

表 5 Weibull 分布参数的区间估计

Tab. 5 Point estimation and interval estimation of Weibull distribution parameters

安全行车周期样本 Driving safety cycle sample	形状参数 Shape parameter $\hat{m}$	尺度参数 Scale parameter $\hat{\eta}/a$	位置参数 Location parameter $\hat{\gamma}/a$	决定系数 Coefficient of determination $r^2/\%$
$\{\hat{t}_{(i)}\}$	1.11	153.61	133.00	97.05
$\{\hat{t}_{(iL)}\}$	1.07	119.29	128.00	96.62
$\{\hat{t}_{(iU)}\}$	1.19	202.43	140.00	96.28

表 6 安全行车周期特征信息汇总

Tab. 6 Summary of safe driving cycle characteristic information

既定可靠度 Established reliability $R(t)$	安全行车可靠信息区间估计 Reliable information for driving safety interval estimation t_R/a	平均安全行车周期区间估计 Mean driving safety cycle interval estimation $E(T)/a$
$R = 0.98$	[131, 148]	[244, 331]
$R = 0.95$	[135, 156]	
$R = 0.90$	[142, 170]	
$R = 0.80$	[157, 197]	

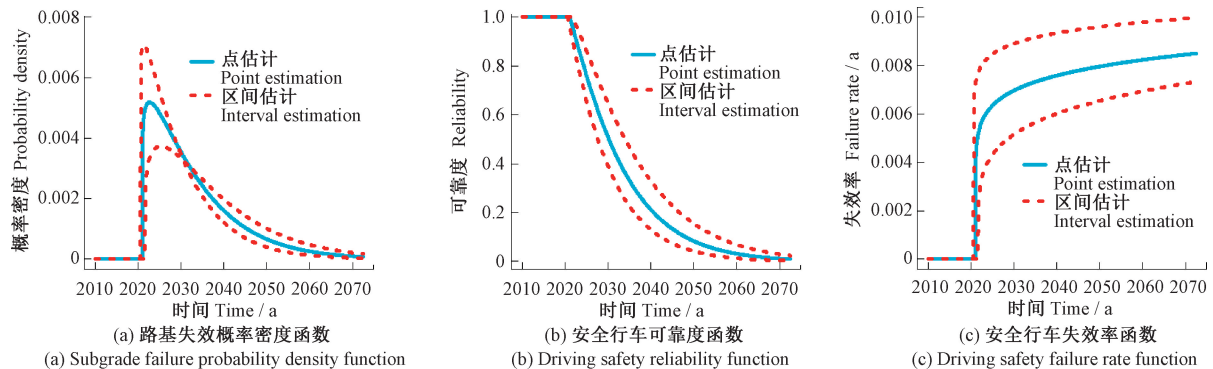


图 2 冻土区安全行车可靠性指标

Fig. 2 Reliability index of driving safety in the permafrost area

格尔木至拉萨 K1449+350~K1450+300 路段路基类型均为一般路基、骨架护坡,自然病害融沉现象频发,冻土区安全行车可靠性指标图像表明:

1) 安全行车可靠度 $R(t)$ 持续到 2020 年 9 月至 2021 年 9 月区间保持在高水平;短期预测表明,安全行车可靠度将在 2020 年 12 月至 2022 年 5 月区间下

降至安全阈值 $R = 0.98$; 长期预测表明将在 2062 年 4 月至 2072 年 6 月区间安全行车可靠度趋于 0, $R < 0.05$, 可靠度随时间递减趋势符合预期。

2) 安全行车失效率在 2010 年 1 月至 2021 年 2 月时间段内几乎为 0; 短期预测表明, 安全行车失效率将在接下来的两年时间骤升至 5.90×10^{-3} 年; 长期预测表明, 从 2023 年 2 月到 2072 年 6 月, 安全行车失效率将缓慢攀升至 8.49×10^{-3} 年左右, 即路基发生失效的概率不断升高。

3) 路基平均安全行车周期 $E(T)$ 运算结果显示从 2010 年 1 月开始测算, 格尔木至拉萨 K1449+350 ~ K1450+300 路段路基平均安全行车周期截止在 2030 年 5 月至 2037 年 8 月区间; 按照既定可靠度水平 $R = 0.95$ 为例, 安全行车周期延续到 2022 年 4 月至 2023 年 1 月区间。

基于可靠性指标完成评估及其数据处理结果, 在工程上建议: 多年冻土路基的维修养护, 应以保护路基系统热力学、力学平衡状态为原则, 避免破坏路基系统内的自然环境状态, 从而保证多年冻土区路基长期稳定, 不宜在暖季后期增加路基荷重, 避免在路肩和顶堆载, 应在日常及时完成多年冻土路基的检查。

4 结论

通过以上分析研究, 得出以下结论:

1) 基于青藏铁路冻土区路基高程沉降数据完成高原冻土区行车安全可靠评估, 首先通过 Sarima 模型纳入季节性影响, 充分考量冻胀融沉两大病害现象, 所建立的路基类型沉降退化轨迹预测结果 $\hat{y}_i(k)$ 长期有效, 且被置信化区间覆盖; 其次根据既定阈值 D_t 完成路基失效时间 T 的外推, 运用最小二乘法实现 Weibull 分布参数估计; 最后完成安全行车可靠度、失效率及平均安全行车周期和可靠信息推断。

2) 着重考察位于格尔木至拉萨 K1449+350 ~ K1450+300 路段路基类型均为一般路基、骨架护坡的行车安全, 具体实现路基失效概率密度函数 $f(t)$ 及累积分布函数 $F(t)$ 和安全行车可靠度函数 $R(t)$ 及失效率函数 $\lambda(t)$ 的呈现, 明确平均安全行车周期 $E(T)$ 及可靠信息 t_R 的推导, 并适宜给出工程建议。

3) 对于路基变形甚微或融沉现象较为轻微路段的可靠性评估, 本文建立的安全行车预警模型可能会预测其中长期内行车风险几乎为零, 由于预测模型的预测步长有限, 涉及定时截尾数据的复杂处理, 未来可结合 Lstm 神经网络与 Sarima 模型完成冻土区路基可靠性评估。

参考文献 (References)

[1] 高宝林, 孙志忠, 董添春, 等. 青藏铁路路基下融化夹层特征及

其对路基沉降变形的影响[J]. 冰川冻土, 2015, 37(1): 126-131.

GAO BaoLin, SUN ZhiZhong, DONG TianChun, et al. Characteristics of thawed interlayer beneath embankment of the Qinghai-Tibet Railway in permafrost regions and its effect on embankment settlement deformation[J]. Journal of Glaciology and Geocryology, 2015, 37(1): 126-131 (In Chinese).

[2] 汪双杰, 金龙, 穆柯, 等. 高原冻土区公路路基病害及工程对策[J]. 中国工程科学, 2017, 19(6): 140-146.

WANG ShuangJie, JIN Long, MU Ke, et al. Distresses and countermeasures of highway subgrade in plateau permafrost regions[J]. Strategic Study of CAE, 2017, 19(6): 140-146 (In Chinese).

[3] 王进昌, 吴青柏. 青藏铁路冻土区路桥过渡段沉降原因分析[J]. 冰川冻土, 2017, 39(1): 79-85.

WANG JinChang, WU QingBai. Settlement analysis of embankment-bridge transition section in the permafrost regions of Qinghai-Tibet railway[J]. Journal of Glaciology and Geocryology, 2017, 39(1): 79-85 (In Chinese).

[4] 洪兆阳, 金双根. 利用时序 PS-InSAR 监测青藏高原冻土区地表形变[J]. 测绘通报, 2021(1): 35-40.

HONG ZhaoYang, JIN ShuangGen. Permafrost deformation in Qinghai-Tibet plateau time-series PS-InSAR[J]. Bulletin of Surveying and Mapping, 2021(1): 35-40 (In Chinese).

[5] 余雨薇, 陈卫雄, 郭继林, 等. 温度与振动耦合作用下青藏铁路路基沉降可靠性分析[J]. 铁道科学与工程学报, 2022, 19(2): 376-383.

YU YuWei, CHEN WeiXiong, GUO JiLin, et al. Reliability analysis of subgrade settlement of Qinghai-Tibet railway under the coupling effect of temperature and vibration[J]. Journal of Railway Science and Engineering, 2022, 19(2): 376-383 (In Chinese).

[6] LIU J X, ZHOU Z J, ZHONG Y C, et al. Prediction of the dissolved gas concentration in power transformer oil based on SARIMA model[J]. Energy Reports, 2022, 8(1): 1360-1367.

[7] PEIRANO R, KRISTJANPOLLER W, MINUTOLO M C. Forecasting inflation in Latin American countries using a SARIMA-LSTM combination[J]. Soft Computing, 2021, 25(16): 10851-10862.

[8] FARSI M, HOSAHALLI D, MANJUNATHA B R, et al. Parallel genetic algorithms for optimizing the SARIMA model for better forecasting of the NCDC weather data[J]. Alexandria Engineering Journal, 2021, 60(1): 1299-1316.

[9] JOSEF A, PETER T. Automatic SARIMA modeling and forecast accuracy[J]. Communications in Statistics-Simulation and Computation, 2019, 50(10): 1-22.

[10] GUO J Y, LI Y F, PENG W W, et al. Bayesian information fusion method for reliability analysis with failure-time data and degradation data[J]. Quality and Reliability Engineering International, 2022, 38(4): 1944-1956.

[11] DUAN C Q, DENG C, LI N. Reliability assessment for CNC equipment based on degradation data[J]. The International Journal of Advanced Manufacturing Technology, 2019, 100(1): 421-434.

[12] PALAYANGODA L K, BUTLER R W, NGH K T, et al. Evaluation of mean-time-to-failure based on nonlinear degradation data with applications[J]. IIE Transactions, 2021: 1-18.

- [13] 殷泽凯, 唐家银, 何平. 恒加双应力型相对熵信息融合可靠性评估模型[J]. 机械强度, 2021, 43(4): 841-848.
YIN ZeKai, TANG JiaYin, HE Ping. Reliability assessment model based on relative entropy information fusion with double constant accelerating stresses[J]. Journal of Mechanical Strength, 2021, 43(4): 841-848(In Chinese).
- [14] 何晓群, 刘文卿. 应用回归分析[M]. 4版. 北京: 中国人民大学出版社, 2015: 19-36.
HE XiaoQun, LIU WenQing. Applied regression analysis[M]. 4th ed. Beijing: China Renmin University Press, 2015: 19-36(In Chinese).
- [15] 赵宇. 可靠性数据分析[M]. 北京: 国防工业出版社, 2011: 80-83.
ZHAO Yu. Data analysis of reliability[M]. Beijing: National Defense Industry Press, 2011: 80-83(In Chinese).
- [16] 冯钧, 刘伟, 谭龙, 等. 高维 Copula 函数的动态系统可靠性模型研究[J]. 机械强度, 2022, 44(1): 86-94.
FENG Jun, LIU Wei, TAN Long, et al. Research on dynamic system reliability model of high dimensional copula function[J]. Journal of Mechanical Strength, 2022, 44(1): 86-94(In Chinese).
- [17] 王帅, 高鹏, 刘畅, 等. 基于时域-误差模型的运动时变可靠性研究[J]. 机械强度, 2020, 42(4): 869-874.
WANG Shuai, GAO Peng, LIU Chang, et al. Research on time-dependent kinematic reliability based on time domain error model[J]. Journal of Mechanical Strength, 2020, 42(4): 869-874(In Chinese).
- [18] WANG Y F, HUANG Y F, LIAO W C. Degradation analysis on trend gamma process[J]. Quality and Reliability Engineering International, 2022, 38(2): 941-956.