基于图的数据驱动发现：高压输电线路电晕诱发噪声与无线电干扰的可解释定律

Hao Xu[1,2], Yuntian Chen[1,3,*], Chongqing Kang[2,*] 和 Dongxiao Zhang[1,4,*] 浙江省工业智能与数字孪生重点实验室，东方理工大学，中国浙江宁波 315200 [2] 清华大学电气工程系，中国北京 100084 [3] 宁波数字孪生研究院，东方理工大学，中国浙江宁波 315200 [4] 岭南大学高级研究院，中国香港屯门 * 通讯作者。

电子邮箱：ychen@eitech.edu.cn（Y. Chen）；cqkang@tsinghua.edu.cn（C. Kang）；dzhang@eitech.edu.cn（D. Zhang）。

摘要

全球向可再生能源转型，亟需发展特高压交流输电，以弥合远程能源与城市需求之间的差距。尽管特高压电网具有更大的容量和更高的效率，但其建设往往受到电晕诱发的可听噪声（AN）和无线电干扰（RI）的制约。由于这些排放必须符合严格的环境合规标准，因此，准确预测对于特高压基础设施的大规模部署至关重要。现有工程实践通常依赖经验定律，其固定的对数线性结构限制了准确性和外推能力。本文提出一种单调性约束图符号发现框架 Mono-GraphMD，用于发现描述电晕诱发的可听噪声（AN）和无线电干扰（RI）的简洁、可解释定律。该框架揭示了表面梯度、分裂数和直径之间的非线性相互作用如何决定高场强排放，并能够对电晕笼数据以及来自多个国家、采用多达16分裂导线的实际特高压线路进行准确预测。与黑箱模型不同，所发现的闭式定律具有很强的可迁移性和可解释性，可用于多种场景下的快速预测，从而助力工程设计。

关键词： 方程发现、电晕放电、高压输电、可听噪声、无线电干扰

引言

高压输电走廊已成为不断演变的全球能源格局中不可或缺的组成部分。远距离可再生能源资源的大规模并网，以及发电中心与需求中心之间的地理错配，都要求通过数百乃至数千公里的距离输送大容量电力（图 1a）。[1] 此外，输电线路对环境的影响正受到越来越多的关注。一个主要问题是电晕效应，它会产生可听噪声（AN）[2] 和无线电干扰（RI）[3]。这些排放会直接影响当地的声学与电磁环境[4], ，引发公众接受度方面的担忧，并促使许多司法管辖区实施严格的监管限值（图 1a）。

电晕放电的物理机制在定性层面已得到理解，涉及空气分子的电离和空间电荷的形成。[5] 然而，要快速、准确地定量预测电晕引起的 AN 和无线电干扰（RI），仍然面临挑战。物理机制之间错综复杂的耦合，阻碍了第一性原理模型的发展，使其无法通过解析方式将这些现象与导线几何结构、表面电场等基本设计参数联系起来。因此，输电设计实践历来依赖经验推导公式（图 1b）。[6] 邦纳维尔电力管理局（BPA）等机构的开创性工作[7] 主要基于受控电晕笼实验和线路测量，建立了相关关系。[8] 这些方法主要采用经验公式，将表面电场、导线直径、分裂数和观测距离与预测值之间的关系表示为对数函数，并根据特定运行条件下的数据拟合系数。[9,10] 尽管这些表达式计算效率高，但其简化的对数线性形式本质上无法捕捉潜在的非线性和变量间的相互依赖关系[11]。因此，这一局限性会降低预测准确性，并限制模型在更广泛的导线配置和运行条件下的泛化能力。近年来，机器学习（ML）已成为预测任务的有力替代方案。相关向量机[12] 和线性混合效应模型[13] 等技术已用于改进 AN 和 RI 预测，而主成分分析与神经网络相结合，则被用于基于相关特征构建预测模型[14]。尽管这些数据驱动模型的预测性能可能优于传统经验公式，但它们主要以“黑箱”方式运行，内部逻辑不透明，无法提供直观的闭式数学表达式来阐明输入参数与电晕放射之间的潜在关系。[15] 黑箱模型固有的可解释性不足，会削弱其在关键决策中的可靠性，并妨碍其纳入既有设计标准（图 1b）。此外，这些模型往往与现有物理分析框架相互割裂，而解析方程则可以无缝

地与理论体系相协调，从而支持更广泛的比较研究。从计算角度看，高维数据驱动算法会带来显著的计算开销，因此不适用于资源受限的边缘计算场景。相比之下，参数较少且计算需求极低的显式数学公式，非常适合低功耗、实时工程应用。因此，亟须将复杂的高维映射转化为简洁的数学表达式。这样的形式化处理能够弥合先进数据分析与实际工程需求之间的差距，促进发现兼具计算效率与物理一致性的符号知识。

与此同时，数据驱动方程发现[16,17]，这一旨在直接从观测数据中提炼出明确且易于人类理解的数学表达式的领域，正受到越来越多的关注。符号回归[18]是一种代表性技术，它在各种数学形式的空间中进行搜索，以找出最符合数据的方程，而无需预先设定模型结构。通过将基于神经网络的搜索与受物理学启发的启发式方法相结合，AI Feynman[19]和 uDSR[20]等先进算法已被证明能够直接从观测数据中重新发现已知物理定律并揭示新的定律。微分方程发现领域的相关研究也进一步展现了识别动力系统模型[21,22]的潜力。这些进展凸显了人工智能驱动的方程发现超越单纯经验曲线拟合、带来真正科学洞见的变革性潜力。尽管人工智能驱动的方程发现已在多个科学领域[23,24]，取得成功，但将其应用于复杂工程问题时仍面临障碍。这些挑战主要源于高维度，以及来自第一性原理的约束相对薄弱且间接。例如，Yang 和 Park[25]采用进化计算，为预测暖通空调管线产生的可听噪声推导新公式。然而，该方法固有的局限性限制了可发现项的复杂程度，最终未能识别出对于准确预测至关重要的、更合适的非线性相互作用。大多数符号回归算法通过将算子和潜在变量组合成表达式树[26]来探索解空间。然而，随着变量和允许的相互作用项数量增加，可能表达式的组合爆炸会导致严重的维数灾难，损害优化效率与稳定性。缺乏稳健的结构约束时，搜索往往会收敛到过度简化的局部极小值，或过于复杂、难以解释且对数据过拟合的形式。因此，在存在测量噪声或数据覆盖范围有限时，要兼顾简约性与准确性尤其困难。此外，标准目标函数主要关注数值误差，却忽略了定性领域知识，例如预期的单调趋势，或极端条件下的有界响应。因此，所得公式虽然可能达到

nu数值准确性，却违反基本趋势，因而限制了其在工程领域的实际应用价值 .

为解决这些局限性，本文提出了具有单调性的基于图的模型发现框架 Mono-GraphMD（图 1c）。通过将数学表达式表示为计算图，所提出的框架能够更加灵活地施加结构约束，从而提高在由多个变量构成的高维空间中进行搜索的效率和有效性。此外，我们通过引入单调性惩罚，将领域知识直接融入优化目标，确保发现的公式在定义域内遵循已知的物理趋势，并提高其泛化能力和可靠性。将该框架应用于来自中国特高压电晕笼试验设施的数据集后，我们发现了用于预测可听噪声 A计权声压级（L_{AN}）和无线干扰激发函数（RIEF）的新颖且可解释的公式。这些新表达式纳入了以往被忽略的关键变量之间的耦合项。比较分析表明，在项数所体现的复杂度相当的情况下，我们发现的公式较既有经验模型具有更高的预测准确度。通过预测不同国家多个全尺寸特高压交流（AC）项目沿线的可听噪声和无线电干扰水平，进一步验证了这些发现公式的实际有效性；预测结果与现场实际测量值高度吻合。这些结果不仅验证了所发现的电晕放电模型的有效性，也凸显了以领域知识为基础的人工智能驱动的模型发现，在科学与工程领域突破人工推导经验公式局限方面的更广泛潜力。

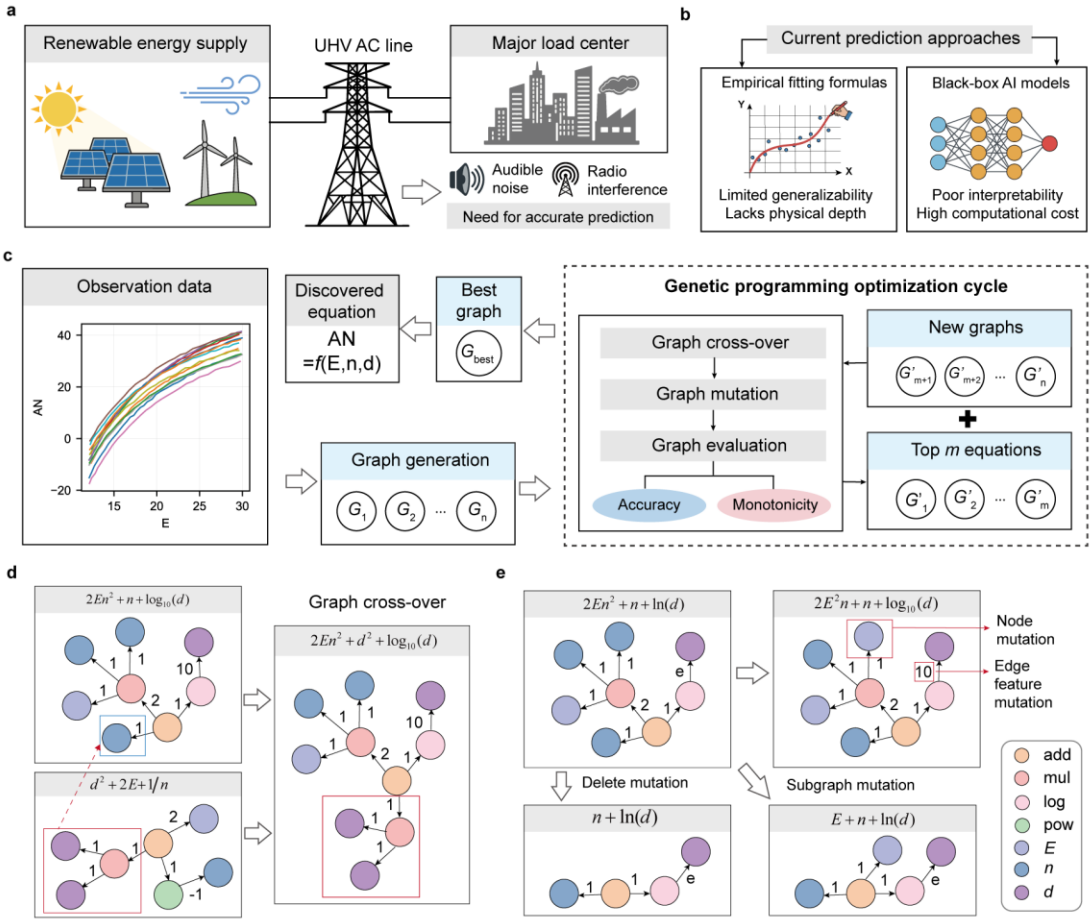


图 1. 所提出的基于图的模型发现框架 Mono-GraphMD 的背景与概述。 （**a**）示意图展示了高压输电线路将可再生能源从发电站输送至主要负荷中心时产生的电晕引起的可听噪声和无线电干扰。（**b**）现有预测方法（包括经验模型和黑箱人工智能方法）的主要局限性总结。（**c**）模型发现过程的流程图，包括图生成、评估和优化。（**d**）图交叉示例。（**e**）图变异示例。此处，方程以图结构表示，其中节点代表变量和算子，边代表计算依赖关系，边特征则编码局部运算和系数。

方法

用于发现物理定律的单调性约束框架

Mono-GraphMD 是一种融入单调性约束的基于图的模型发现框架。其架构由两个紧密耦合的组件构成：由图模板库构建的候选表达式的基于图的表示，以及基于遗传编程、直接在表达式图空间中运行的进化优化过程（图 1c）。

方程的基于图的表示

核心思想是将每个候选方程视为一个有向无环图（DAG）。有向无环图由一组节点和有向边组成。在 Mono-GraphMD 中，节点表示算子（如加法、乘法、幂运算和对数）或变量，而边则编码它们之间的计算依赖关系。每条边都关联一个边特征，用于参数化局部运算。例如，对于幂运算节点（幂运算），其入边上的边特征编码指数；对于对数节点（对数），边特征指定底数；对于其他算子，边特征则作为标量系数，用于乘以相应子图的贡献。通过这种方式，表达式的结构由图的拓扑结构和标签来表示。

所提出的图表示采用 n 元连接模型和属性边框架，不同于传统的二叉表达式树。二叉树受严格的二分结构限制，运算符节点最多只能有两个子节点；而我们的架构允许加法节点和乘法节点连接任意数量的子图。关键的是，这一转变显著降低了拓扑复杂度：许多原本需要表示为额外节点的量，例如系数或运算符特定参数，改为编码为附加在计算关系上的边属性。这种重新参数化减少了显式节点的数量和拓扑深度，同时将操作数子图与关系特定属性分离。因此，运算符节点的子节点不再是同质的：操作数子图表示组合结构，而边属性则表示与运算符相关联的参数，这些参数与父运算符一同计算。灵活性的提升使表达式能够以紧凑形式表示；对于涉及众多变量和多重非线性相互作用的问题，这一优势尤为明显，因为基于树的编码往往会变得不必要地深且不平衡。

由于不受约束的图结构具有极高的结构自由度，直接探索完整图空间在计算上代价高昂，而且容易生成无效或缺乏信息的表达式。为解决这一问题，
Mono-GraphMD 采用图模板机制，其核心是一个加法根节点，用于叠加不同的函数分量。这种加法结构使模型能够捕捉全局缩放以及独立特征的累积效应。在这一框架中，三种主要函数原型以随机方式整合：多项式子图、有理式子图和对数变换。多项式模块通过幂的乘积形式实现，旨在明确刻画多变量交互，使系统能够捕捉输入特征之间的非线性依赖关系。有理式子图则引入

倒数关系和归一化效应，而对数模板则支持尺度不变变换，并能够处理指数数据分布。通过将搜索空间约束在这些具有物理意义的骨架上，生成的计算图能够与可解释的数学形式直接对应。这种结构化初始化确保发现的方程不仅仅是数值拟合结果，还由函数构件组成，为后续的符号可解释性和领域特定分析奠定基础。

将物理单调性融入方程发现

对于任意固定的图结构，通过对观测数据进行回归来估计与各项相关的数值系数，并计算决定系数（R^2）以评估表达式的数值准确性，该表达式记为L_{acc}=1-R^2。此外，Mono-GraphMD 纳入了单调性约束，以编码关于目标变量对各变量定性依赖关系的先验知识。对于给定变量x_j，指定一个符号$s_j \in [+1,-1]$ ，表示预期函数在指定定义域D_j. 内整体呈非递减（s_j=+1）或非递增（$s_j = -1$）趋势。对一组具有代表性的有限取值$\{x_j^{(l)}\}_{l=1}^{L} \subset D_j$ 进行采样，并将所有其他变量固定为一组标称值。变量j的单调性损失定义如下：

$$L_{mono,j}=\sum_{l=1}^{L-1} \max\left(0, -s_j[\hat{y}(x_j^{l+1}) - \hat{y}(x_j^{l})]^2\right) \tag{1}$$

该损失会对违反指定单调趋势的情况施加惩罚。因此，总单调性惩罚为$L_{\text{mono}}=\sum L_{mono,j}$。引导进化搜索的总体损失定义如下：

$$L=L_{\text{acc}}+\lambda_{\text{mono}}L_{\text{mono}} \tag{2}$$

其中，$\lambda_{\text{mono}}>0$ 是权重参数。L的值越低，表示方程既能很好地拟合数据，又能在超出观测数据范围的定义域内遵循单调性先验。在 Mono-GraphMD 中，稀疏性通过结构而非显式正则化来控制：在图层级设定项数上限，从而确保发现的方程保持紧凑，无需额外的稀疏性惩罚。

图空间中的遗传编程优化

基于图的表示下，Mono-GraphMD 使用遗传编程方案进行结构优化（图 1c）。通过从模板中采样，并将实例化后的模板组合成多项表达式，生成初始方程图种群。随后，通过反复执行交叉、变异和评估来演化该初始种群。

交叉在子图层面进行。对于两个父图，连接到

顶层加法节点（即各个独立项）的子图会被随机选取并交换（图 1d）。用代数术语来说，这相当于在两个候选方程之间交换一个或多个项。由于每一项本身都是基于有效模板构建的，因此后代仍属于允许的结构类别。变异通过三种互补方式进行（图 1e）：（i）边特征变异，即扰动选定的边特征；（ii）子图变异，即将现有子图（项）替换为新采样的模板实例化；（iii）子图添加/移除，即向连接到顶层加法节点的集合中添加或移除子图，但项数不得超过规定的上限。重要的是，所有变异都通过相同的模板机制执行，从而确保变异后的图保持有效，且结构复杂度不会失控。完成交叉和变异后，根据结合数据拟合精度与单调性正则化的损失函数对种群进行评估和排序。排名前 50% 的图将作为父代保留到下一代，其余图则由新生成的随机图取代，以保持搜索的多样性。这个进化循环会持续迭代，直至达到规定的代数，并将损失最低的图选作最终发现的方程。

结果

交流输电线路可听噪声预测公式的发现

在交流（AC）激励下，当瞬时表面电场超过空气击穿强度时，输电导线便会发生电晕放电。由此产生的空间电荷被限制在导线附近的薄层内，因为交流电场的极性快速反转会周期性地扫除电荷，从而阻止大范围空间电荷云的形成[27]。这意味着局部电晕现象，包括功率损耗、无线电干扰和可听噪声，主要由局部电气和几何条件决定，几乎不受完整三维线路配置的影响。[28] 这一特性使得可以通过局部测试装置（如单相试验线路或电晕笼）来表征电晕性能。因此，根据此类实验得出的经验关系，可以可靠地外推，用于预测具有任意几何配置的输电线路的表现。

用于方程发现的可听噪声数据来自针对十三种分裂导线配置开展的电晕笼实验。[29] 测试矩阵涵盖由四种子导线类型（LGJ630/45、LGJ500/35、LGJ400/35 和 LGJ300/50）组成的8、6和4子导线束。对于每种配置，均以不同电压等级对导线束施加电压，以实现宽范围的表面电场梯度。精密噪声

分析仪记录发出的声音，并将其转换为A计权声压级L_{AN}，其参考值为μW/m，定义如下：

$$L_{AN}^{(\mu W/m)}=10\log_{10}\left(\frac{P_{AN}}{1\mu W/m}\right) \tag{3}$$

其中，P_{AN} 为测得的单位长度A计权声功率。在部分文献中，使用以 pW/m 为参考值的L_{AN}，两者之间的换算关系如下：

$$L_{AN}^{(\mu W/m)}=L_{AN}^{(pW/m)}-60 \tag{4}$$

根据既有证据[29]，在标准设计范围内，束间距对可听噪声的影响很小，因此排除该参数，以避免不必要的模型复杂度。用于方程发现的输入变量因此为表面梯度（E，千伏/厘米）、子导线数量（n）和子导线直径（d，厘米），目标变量为A计权声压级（L_{AN}）。

基于图的方程发现框架采用了一个算子集（+, ×, *pow*, *log*）、500 的种群规模和 200 次优化迭代。尽管该算子集规模较小，但其函数表达能力较强，因为减法可以通过带负乘法系数的加法实现，而除法则表示为乘以逆幂。目标函数中纳入了权重为 0.01 的单调性惩罚。对于当前应用，预期L_{AN} 在与设计相关的范围内随E、n和d单调递增。通过对这三个变量分别施加全局单调性约束，将这些预期编码进模型。为考察准确性与结构简洁性之间的权衡，将候选公式的最大项数限制为 3、4 和 5（图 2a）。在这些设置下得到的最佳三项、四项和五项表达式如下：

$$L_{AN}=0.0878En+72.3\log(d)-\frac{648.7}{E\log(E)} \tag{5}$$

$$L_{AN}=0.093En+55.02\log(d)-\frac{591}{Ed^2}-\frac{5448}{E^2} \tag{6}$$

$$L_{AN}=0.0116n^2d-\frac{102.4n}{E\ln(E)+d^2}+9.216d+19.13\ln(n)-\frac{677.3}{E} \tag{7}$$

为了进行比较，我们还对同一数据拟合了一个常规多项式型回归模型，其结构与经典公式相似：

$$L_{AN}=1.022n+10.4d+30.839-\frac{933.633}{E} \tag{8}$$

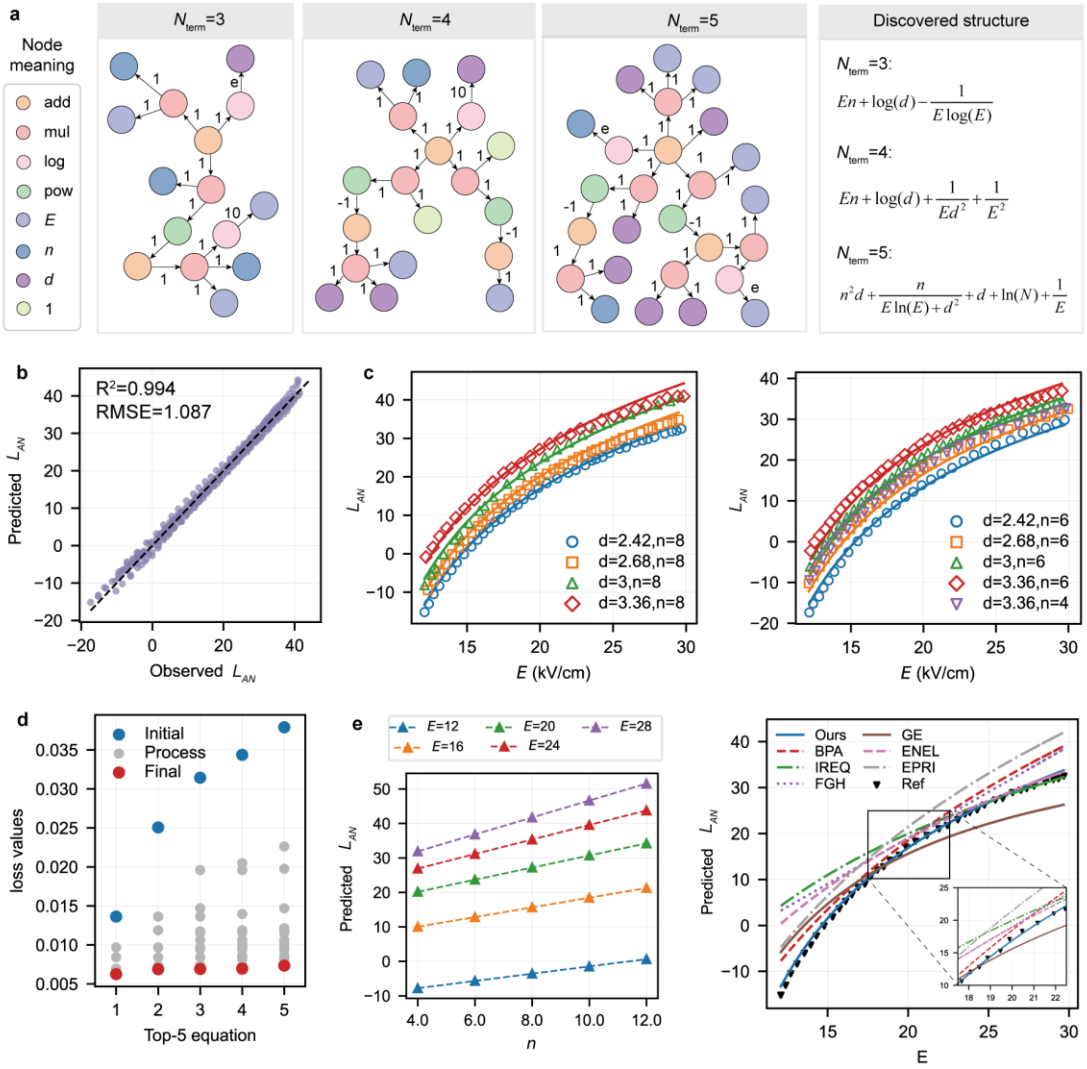


图2. A计权声压级L_{AN}的数据驱动方程发现结果。（**a**）不同最大项数下发现的最优图结构和公式结构。（**b**）所有数据中预测值与观测到的L_{AN}之间的比较。（**c**）对于子导线根数L_{AN}为n=8、6和4且子导线直径d不同的情况，预测值与观测到的值之间的比较。其中，散点表示观测数据，曲线表示发现的公式给出的预测值。（**d**）优化过程中前5个方程的损失值。（**e**）不同子导线根数L_{AN}、n和表面梯度E下的预测值。（**f**）发现的公式与8×LGJ300/50分裂导线的经典经验模型之间的比较。

在整个数据集上，三个发现的公式的均方根误差（RMSE）分别为 1.087、1.033 和 0.917，平均相对误差（MRE）分别为 0.297、0.308 和 0.248，对应公式（5）至（7）；而公式（8）中的多项式基线的均方根误差为 1.714，平均相对误差为 0.336。因此，即使是最简单的发现的三项定律

与经过精心调优的多项式表达式相比，均方根误差降低了约 40%。公式 (5)–(7) 的数学结构揭示了潜在的依赖关系。发现的方程中包含线性和对数线性分量，这与传统经验定律采用的形式相呼应。在公式 (5) 中，非线性有理项 $^{1}{}_{E}\log(E)$ 有效地捕捉了可听噪声在高场梯度下呈现的饱和现象，而使用传统方程发现方法难以识别这一现象。本文选择公式 (5) 进行后续预测和分析。尽管公式 (7) 的均方根误差最低，但与结构更紧凑的公式 (5) 相比，其改进幅度有限，不足以抵消其函数形式增加的复杂性。因此，公式 (5) 在准确性与可解释性之间实现了最佳平衡，更适用于实际应用。图 2b 和 c 展示了预测性能。预测曲线几乎与实验标记点重合，捕捉到了L_{AN} 随E非线性增加的趋势，以及随着d增大，曲线整体系统性上移的现象。这种吻合表明，发现的公式正确捕捉了可听噪声对相关变量的依赖关系。此外，优化轨迹显示，排名前 5 的方程的损失稳步下降，表明基于图的搜索在单调性约束下可靠收敛（图 2d）。图 2e 清楚显示了发现的定律具有单调性：在整个定义域内，预测曲线随变量E平滑上升，且没有局部下降或非物理振荡。

所发现的可听噪声预测公式与现有经验模型的比较

随后，利用电晕笼实验数据以及来自特高压交流输电线路的现场测量数据，将发现的公式与经典经验模型进行性能基准比较。对于具有代表性的导线配置，比较了发现的公式与广泛使用的经验公式（包括 BPA、IREQ、FGH、GE、ENEL 和 EPRI）所预测的可听噪声水平L_{AN} 随表面梯度E 的变化（图 2f）。这些公式的详细信息见表 S1。图中的黑色三角形表示通过侧向电容器测量获得的参考值。可以观察到，在整个E范围内，发现的定律对应的曲线都紧密贴合参考点，仅在表面梯度较低和较高时存在小幅偏差。在低场区域，几乎所有经验模型都系统性地高估了测量值。在E的中间范围内，若干经验模型与数据吻合良好，尤其是 BPA 模型。不过，大多数模型相对于参考值仍存在偏差。在较高的表面梯度下，大多数经验公式（如 FGH 和 BPA）给出的增幅明显更陡

高于观测到的增幅，导致可听噪声被显著高估。相比之下，ENEL 和 IREQ 模型与观测数据吻合良好。总体而言，现有经验模型仅在有限的表面梯度范围内具有良好的预测性能，而发现的定律在所考察的整个范围内始终保持准确。此外，经验模型通常包含四个或更多项，并且受限于特定的对数线性结构。相比之下，发现的公式仅使用三项，却在整个实验范围内优于经典经验定律。这些发现表明，与仅凭人类直觉进行经验建模相比，数据驱动方程发现更有利于找到简洁而准确的公式。

表1. 全尺寸特高压输电线路可听噪声预测公式与现有经验模型的比较。距离的单位为米，测量噪声和预测噪声的单位为分贝。预测值与观测值之间的差异列于括号中。粗体条目表示各项比较中的最佳表现。

线路	距离	实测	BPA	ENEL	EPRI	Ours
工程_特高压_8	15	50.2	51.7 (+1.5)	51.7 (+1.5)	**51.0 (+0.8)**	52.1 (+1.9)
项目_特高压_12	15	55.8	53.6 (-2.2)	52.9 (-2.9)	48.2 (-7.6)	**53.7 (-2.1)**
项目_特高压_16	15	50.8	51.9 (+1.1)	51.6 (+0.8)	43.2 (-7.6)	**51.4 (+0.6)**
HN-JZ	15	45.2	46.7 (+1.5)	47.4 (+2.2)	46.0 (+0.8)	**46.0 (+0.8)**
HB-ZX	10	46.7	47.2 (+0.5)	47.9 (+2.2)	46.5 (-0.2)	**46.5 (-0.2)**
	20	46.9	47.8 (+0.9)	48.4 (+1.5)	47.1 (+0.2)	**47.1 (+0.2)**
HB-ZY	10	47.8	**47.9 (+0.1)**	48.6 (+0.8)	47.1 (-0.7)	47.3 (-0.5)
	20	47.0	**47.3 (+0.3)**	49.0 (+1.0)	46.5 (-0.5)	46.6 (-0.4)
平均绝对误差			1.01	1.61	2.30	**0.84**

我们还将发现的方程与成熟的符号回归基线方法进行了比较，包括 PySR[30] 和深度符号优化（DSO）[20]。这些方法确定的函数形式列于表 S1。PySR 得出的表达式相对简单，非线性有限，其结构与传统经验公式较为接近。即使

项数多于我们的模型（四项），PySR 的预测性能仍不及我们的模型（均方根误差=1.152，平均相对误差=0.362）。相比之下，DSO 得出的表达式复杂得多，包含五项，且采用引入更强非线性的有理结构。尽管如此，其预测准确度仍然较低（均方根误差=2.112，平均相对误差=0.540）。总体而言，与这些基线方法相比，我们提出的方程在简约性与预测准确度之间实现了更佳的权衡。

随后，我们进一步在电晕笼环境之外的全尺寸特高压输电线路上评估所发现的可听噪声预测公式的性能。对于每条线路，首先利用该定律计算与当地分裂导线配置和表面梯度相关的 A计权可听噪声级。随后，根据实际线路几何结构，采用标准传播和配置修正，将这些级别换算为指定观测距离处的地面声压级。从导线束级发射到现场测量噪声的完整转换过程详见补充信息 S1.1。

表1对比评估了中国和美国特高压输电线路测得的可听噪声级，以及三种经验模型（BPA、ENEL 和 EPRI）与我们发现的定律给出的预测值。就平均误差而言，我们发现的定律总体误差显著低于经验基准模型，表明在大多数测试条件下，它都取得了最佳性能。对于美国的三个特高压项目，其线路分别采用 8、12 和 16 根子导线组成的导线束；我们的定律在 12 和 16 根子导线的配置下表现最佳。BPA 模型似乎系统性地高估了测量值，这表明其公式存在偏差，反映了人工经验拟合的局限性。相比之下，我们发现的方程在不同案例中的误差并未表现出一致的方向性偏差。值得注意的是，尽管该定律最初是根据仅涵盖 4 至 8 根子导线的电晕笼数据推导而来，但将其用于子导线数量更多的配置时，仍保持了较高的预测准确度，并有效优于成熟的经验模型。这表明该定律在原始训练域之外具有很强的泛化能力。

随后，我们利用中国1000千伏晋东南—南阳—荆门特高压交流输电线路的现场数据进行比较，测量数据采集自焦作（HN-JZ）、钟祥（HB-ZX）和枣阳（HB-ZY）站点，距离为10–20米。发现的定律在HN-JZ和HB-ZX站点均优于现有方程。值得注意的是，尽管经验模型在某些特定情况下（例如HB-ZY案例中的BPA模型）偶尔能取得更高的准确性，但在其他情形下，其预测误差更大，导致总体误差更高。这种不一致性凸显了传统经验模型的固有局限，即其

有效性通常仅限于少数场景。相比之下，发现的定律在各种不同场景下均能给出稳健且一致的预测。

交流线路无线电干扰激发函数的发现

高压交流线路上的电晕放电会产生无线电干扰，表现为射频频段的脉冲电磁噪声，并可能严重影响附近的通信系统。在工程实践中，给定观测点处的辐射干扰水平通常以无线干扰电压或场强表示；其不仅取决于导线固有的电晕活动，还受线路几何结构、杆塔配置和传播条件的影响。为区分这些影响，通常采用无线干扰激发函数（RIEF）的概念[31]。RIEF 表征分裂导线在特定电气条件下产生无线电干扰的固有能力，单位为分贝（dB）。已知 RIEF 随局部表面梯度和导线参数变化的关系后，即可通过施加几何和传播修正，计入相位排列、杆塔高度和观测距离的影响，从而求得地面处的辐射干扰[32]。推导详情见补充信息 S1.2。

为进行数据驱动方程发现，将 RIEF 建模为与可听噪声分析中使用的相同三个预测变量的函数：导线束表面梯度 E、导线束中的子导线数量 n，以及子导线直径 d。实验中的分裂导线配置与可听噪声实验中的配置相同。针对每种配置，通过调节施加电压来激发并控制电晕，使表面梯度覆盖较宽范围。在电晕笼试验中，采用侧置电容器法量化 RIEF值。该方法记录通过测量电容器耦合的干扰电压。随后，按照标准化程序对该电压进行处理，以确定等效 RIEF 值 Γ_{RI}，单位为分贝。

在与可听噪声研究类似的设置下，我们将基于图的方程发现框架应用于该数据集。候选解析形式由算子集（+, ×, 幂, 对数）构建。搜索采用种群规模为 500、优化迭代 200 次的设置。为确保物理一致性，我们在优化目标中加入了权重为 0.1 的单调性惩罚。该正则化基于公认的工程原理： Γ_{RI} 必须随表面梯度、导线束尺寸和导线直径单调递增。与可听噪声的情况相同，这一约束是在超出实验范围的可行域内施加的，从而促进符合物理规律的外推。

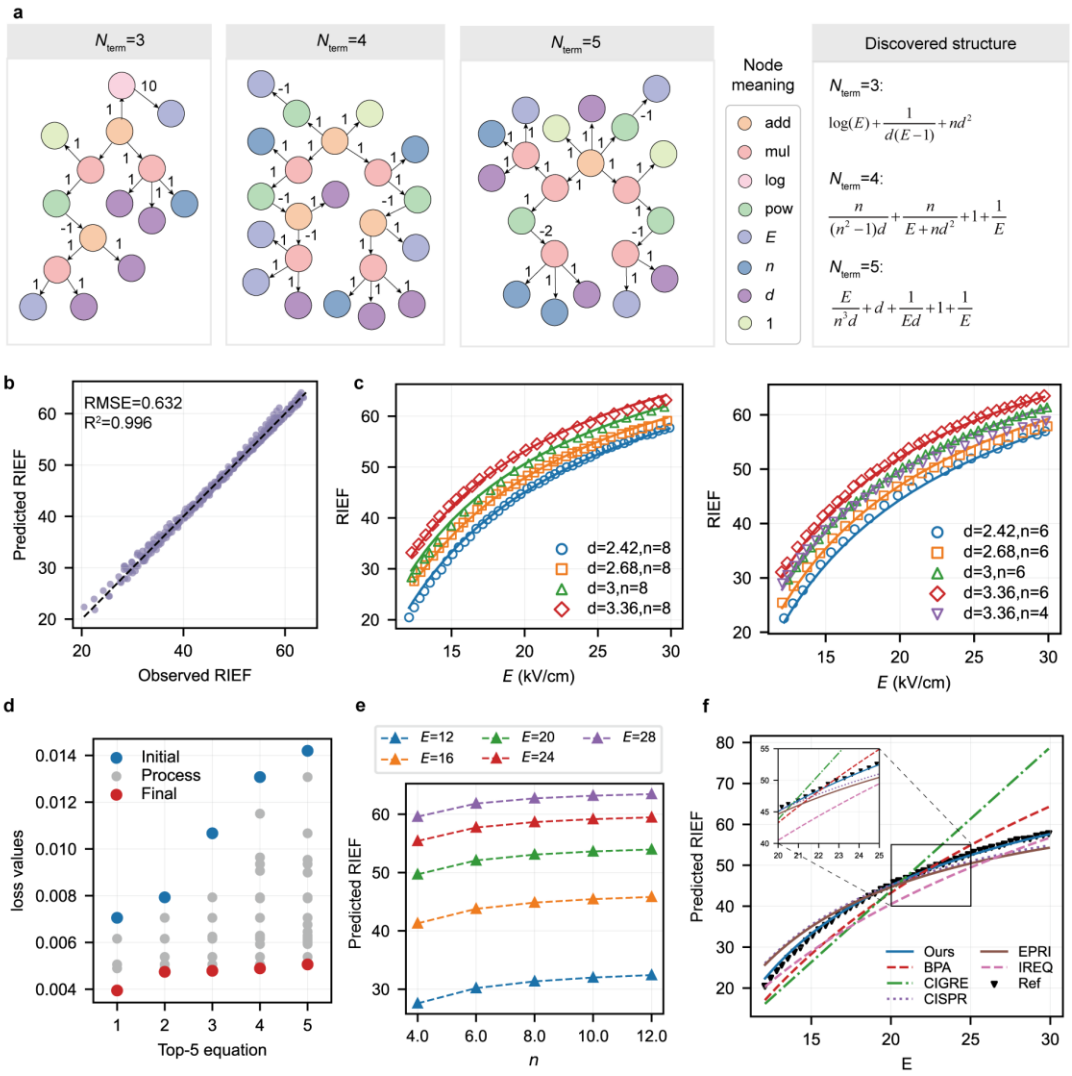


图 3. 无线电干扰激发函数的数据驱动方程发现结果 Γ_{RI}。（**a**）不同最大项数下发现的最佳图结构和公式结构。（**b**）所有数据中预测的 Γ_{RI} 与观测的 Γ_{RI} 之间的比较。（**c**）子导线根数n=8、6 和 4，且子导线直径d不同时，预测的 Γ_{RI} 与观测的 Γ_{RI} 之间的比较。此处，散点表示观测数据，曲线表示由发现的公式得出的预测值。（**d**）优化过程中排名前 5 的方程的损失值。（**e**）不同子导线根数n和表面梯度E下的预测 Γ_{RI} 。（**f**）发现的公式与经典经验模型对8×LGJ300/50 分裂导线的比较。

为探究结构复杂度与预测性能之间的关系，我们将候选表达式中的项数上限设为 3、4 和 5（图 3a）。各项约束下表现最佳的公式如下：

$$\Gamma_{RI}=45.6\log(E)-\frac{819.5}{d(E-1)}+0.07nd^2 \tag{9}$$

$$\Gamma_{\mathrm{RI}}=-\frac{117.2n}{n^2d-d}-\frac{133.5n}{E+nd^2}+98.68-\frac{629.7}{E} \tag{10}$$

$$\Gamma_{\mathrm{RI}}=-\frac{45.87E}{n^3d}+4.499d+72.88-\frac{522.2}{E}-\frac{543.4}{Ed} \tag{11}$$

此外，还拟合了一个受经典经验形式启发的传统多项式型回归，以供比较：

$$\Gamma_{\mathrm{RI}}=6.51d+10.287\log(n)+55.22-\frac{671.7}{E} \tag{12}$$

在完整的电晕笼数据集上，发现的公式（式（9）–（11））的均方根误差分别为 0.931、0.632 和 0.521，相应的平均相对误差分别为 0.017、0.012 和 0.010。与经过仔细调优的多项式基线（式（12），均方根误差 =1.021，平均相对误差=0.019）相比，即使是最简单的发现表达式，均方根误差也降低了约 10%。

式（9）–（11）的代数形式揭示了无线电干扰激励与导线几何结构及表面梯度之间的关系。式（10）和（11）中E的负倒数项捕捉到了观测到的无线电干扰低场抑制现象；在充分发展电晕状态以下，表面梯度的小幅增加会使 RIEF值迅速上升，而在较高梯度下，这种影响趋于饱和，这与 1/E的依赖关系一致。在兼顾准确性、可解释性和易用性的情况下，式（10）的四项定律提供了良好的折衷。与多项式基线相比，该定律的均方根误差降低了约 0.39（图 3b），同时保持了相对简单的函数结构。因此，后续线路级预测采用式（10）作为首选的无线电干扰激励定律。值得注意的是，发现的 RIEF 定律保留了电晕物理学所预期的定性单调趋势（图 3c）。优化过程中采用的单调性惩罚，有效避免了发现的 RIEF 公式随电晕严重程度增加而出现非物理性下降（图 3d）。图 3e 清楚展示了发现定律的单调行为：在整个定义域内，预测曲线均随变量n平滑上升。因此，这些数据驱动表达式不仅数值准确性高，而且其行为也符合关于交流输电线路无线电干扰产生机制的既有认识，为后续预测实际特高压线路的辐射干扰水平提供了可靠依据。

发现的无线电干扰激发函数与现有经验模型的比较

首先，利用生成 RIEF 数据所依据的电晕笼测量结果，对发现的 RIEF 公式的性能进行了评估。图中展示了代表性分裂导线配置下，激励函数随表面梯度 E 变化的预测结果

（图3f）。黑色三角形表示通过侧电容测量获得的参考值，彩色曲线则分别对应发现的定律以及已有的经验公式，包括 BPA、CIGRE、CISPR、EPRI 和 IREQ 表达式。这些经验模型的详细信息见表 S2。在表面梯度约为 12.5 至 30 千伏/厘米的整个范围内，发现的定律对应的曲线与参考数据高度吻合，趋势几乎平行，绝对偏差也极小。在低表面梯度下，EPRI 和 CISPR 等经验模型的误差相对较小，但其预测值仍不及发现的定律。相比之下，CIGRE 和 BPA 公式往往会系统性地低估测量值。有趣的是，随着表面梯度增至中等范围（约 20 千伏/厘米），几乎所有经验公式都能给出较为准确的预测。然而，在表面梯度较高时，这些公式的表现开始出现分歧。在这一范围内，CISPR 和 IREQ 与测量结果仍较为吻合，而 CIGRE 等其他经验模型的误差则明显较大。相比之下，发现的模型在整个表面梯度范围内始终能够准确预测，表现出更优的稳定性和泛化能力。

同样，我们将发现的方程与成熟的符号回归基准方法进行了比较，包括 PySR 和 DSO。这些方法得出的函数形式列于表 S2。对于这项任务，PySR 得出了一个非常复杂的表达式，其中包含嵌套对数以捕捉非线性，以及项之间额外的乘法耦合。尽管其预测性能与我们的方法相近（均方根误差=0.747，平均相对误差=0.014），但所得表达式的简约性显著较差。相比之下，DSO 得出的形式要简单得多，主要由线性分量和基本有理项构成，但其准确性明显更差（均方根误差=3.302，平均相对误差=0.061）。总体而言，这些结果凸显了现有基准方法的一个显著特点：不同问题中发现的表达式复杂度可能相差很大，而且与预测性能并不总是相符。相比之下，我们的框架在不同任务中能更稳定地兼顾简约性与准确性。

除电晕笼实验外，发现的 RIEF 模型还在全尺寸特高压输电线路上得到了验证。在线路级评估中，RIEF 作为内禀源参数，随后沿线路传播，以获得地面观测点处的无线电干扰电平。传播模型将每个微分线路段视为一个等效噪声源，并以其局部激励函数 Γ_n 为特征。各线路段对接收点处场强的贡献受频率相关的传递函数和衰减因子调制，而总干扰则通过对所有线路段和模式进行叠加得到。该变换的详情见补充信息 S1.2。

采用这一流程，将发现的定律所给出的预测值与三种

广泛使用的经验公式（BPA、HVTRC 和 CISPR）进行比较，所用测量数据来自位于美国、加拿大和中国的四条特高压交流线路。表2汇总了测得的无线电干扰水平、模型预测值及相应误差。总体而言，对于不同国家的特高压输电线路，发现的定律展现出稳健的无线电干扰预测准确度，其平均绝对误差显著小于传统经验模型。具体来看，HVTRC公式在北美线路上的表现良好，因为该公式最初是基于类似系统的数据开发的。然而，将其应用于中国特高压线路时，其准确性有所下降，且仍不及我们发现的定律。BPA公式在所有这些场景中均表现出持续的负偏差。值得注意的是，BPA应用于中国特高压线路时偏差显著；这类线路采用更大的分裂导线配置，且运行电压更高。相反，CISPR模型虽然在中国特高压线路上准确性较高，但应用于北美和加拿大场景时，性能则大幅下降。与这些经验模型表现不一的情况不同，发现的定律在不同国家的案例中，表现与最佳区域性模型相当或更优，且不存在系统性偏差。

表2. 全尺寸特高压输电线路所发现的无线电干扰激发函数与现有经验模型的比较。距离的单位为米。预测值与观测值之间的差异列于括号中。粗体条目表示各项比较中的最佳表现。

线路	*n*	测得值	BPA	HVTRC	CISPR	Ours
美国韦恩堡（760 千伏）	4	70.5	69.9 (-0.6)	71.2 (+0.7)	74.4 (+3.9)	**70.1 (-0.4)**
加拿大蒙特利尔（735 千伏）	4	68.0	66.0 (-2.0)	**68.5 (+0.5)**	71.7 (+3.7)	67.3 (-0.8)
中国湖北（1050 千伏）	8	72.0	64.6 (-7.4)	71.5 (-0.5)	71.6 (-0.4)	**72.2 (+0.2)**
安徽，中国（1050 千伏）	8	71.3	63.5 (-7.8)	70.4 (-0.9)	70.5 (-0.8)	**71.1 (-0.2)**
平均绝对误差			4.45	0.65	2.20	**0.40**

综合来看，在电晕笼数据和全尺寸特高压线路上的验证凸显了数据驱动的 RIEF 定律的若干重要特征。首先，该定律能够高保真地再现无线电干扰激励对表面梯度的细致依赖关系，既捕捉到低场起始，也捕捉到高场饱和特性，而传统经验公式只能粗略近似这些特性。其次，当与基于物理原理的

传播模型结合时，同一定律得出的线路级无线电干扰预测值可与现有最佳经验模型相媲美，且在若干情况下表现更优。第三，先前的模型往往存在系统性偏差，这反映出人工构建的方程通常针对特定工况或运行条件类别进行定制，因此难以在该定义域之外良好地泛化。相比之下，发现的定律能够推广到涵盖不同电压等级、分裂导线配置和地理区域的系统。

讨论

本研究表明，高压交流输电线路上的电晕引起的可听噪声和无线电干扰，可以用简洁、数据驱动发现的解析定律来描述。这些定律在保持完全可解释性的同时，优于沿用多年的经验公式。基于对十三种分裂导线配置开展的受控电晕笼实验，本文提出了一种基于图的方程发现框架，用于识别表面电场梯度、分裂导线几何结构与电晕放射的两个关键表征量——A计权可听噪声级和 RIEF——之间的明确关系。发现的公式能够再现不同运行条件下测得的变化趋势，并可成功推广至不同国家的全尺寸特高压线路。

从方法论角度看，本文提出的基于图的方程发现框架为电气工程中的数学建模提供了一条新途径。传统经验方法通常依赖人的直觉来提出候选函数项，再通过简单的线性回归或对数线性回归进行拟合。由此得到的公式通常局限于低阶多项式结构，因此仅在有限的校准范围内有效，其预测准确度和泛化能力都受到限制。尝试通过添加修正项来改进这类经验模型，往往会使表达式变得愈发繁琐，却无法解决模型结构与真实依赖关系之间根本不匹配的问题[33]。与之相反，纯黑箱机器学习模型虽能给出极佳的数值预测，却无法明确揭示电气和几何参数如何影响电晕放射，也不能保证外推结果符合物理规律。在实际应用中，部署这类模型还需要专用的计算环境，与闭式公式的简洁性和可移植性相比并不占优。本文采用的基于图的方程发现框架提供了另一种途径。通过在由小型算子集构建的显式解析表达式中进行搜索，该框架既保留了经验定律的透明性和易实现性，又能直接从数据中推导函数结构，而非依赖临时假设。本研究结果表明，即使将项数所体现的代数复杂度控制在不高于传统

公式的水平，数据驱动发现仍能识别出更贴合测量结果的函数形式，并在超出历史校准范围的配置下保持准确性。

与传统的基于树的符号回归相比，本文提出的基于图的方程发现框架更加灵活，并具备更强大的结构搜索能力[34]。该方法通过图模板生成并优化候选方程，在纳入温和结构先验的同时，保留了多样的潜在函数形式空间。这种方法能够有效抑制不合理的复杂表达式出现，而这类表达式往往困扰着不受约束的搜索。所提出框架的一个至关重要的特点，是将单调性约束显式纳入发现过程。在包括电晕现象在内的诸多工程场景中，输入与输出之间的定性依赖关系都已先验确定。物理原理表明，在运行范围内，可听噪声和无线电干扰应随表面电场增大而单调增加，并且通常也随有效导线尺寸增大而增加。传统的符号回归算法缺乏强制执行这类基本规律的机制，因而经常产生这样的表达式：尽管它们对训练数据的拟合尚可，但在外推时却呈现局部振荡或趋势递减等不符合物理规律的特征。通过在优化过程中惩罚违反既定单调关系的情况，并在扩展的可行域内评估候选方程，我们的框架能够得出既准确拟合观测数据、又与物理趋势全局一致的规律。这为重视可解释性和鲁棒性的工程应用带来了进步。

进一步发展的方向已逐渐显现，同时本研究也存在一些局限。分析仅限于可在电晕笼实验中可靠测量和控制的变量，即表面梯度和基本分裂导线几何结构。因此，发现的定律是在相对有限的环境和表面条件范围内校准的，其在极端气候、复杂地形或严重老化导线条件下的性能尚未得到充分验证。尽管电晕笼配置能够代表典型的特高压分裂导线，但并未涵盖全球在役线路设计的全部多样性，而且迄今为止，现场验证仅涉及数量有限的特高压项目。从这个意义上说，单调性约束和结构先验降低了将结果外推至本研究数据所涵盖范围之外时产生偏差的风险，但并未将其完全消除。此外，方程发现框架本身包含若干可能影响最终公式的超参数，也可能掩盖其他准确性相当但可解释性不同的表达式。总体而言，尽管存在上述限制，结果表明，高压输电设计能够从摆脱固定经验公式中获益良多

转向在明确的结构和物理约束下通过数据驱动发现的解析定律。

致谢

本研究得到国家自然科学基金委员会（批准号：12501744、52288101 和 12572266）、博士后资助计划和中国博士后科学基金会（批准号：BX20250063 和 2024 M761535）、国家重点研发计划（2024YFF1500600）以及宁波甬江人才工程（2022A-242-G）的资助。本研究还得到宁波东方理工大学和宁波数字孪生研究院高性能计算中心的支持。

作者贡献

H. X.、C. K.、Y. C. 和 D. Z. 提出了研究构想，设计了研究方案并分析了结果。H. X. 开发了算法，完成了计算并生成了结果和图表。H. X.、C. K.、Y. C. 和 D. Z. 撰写并编辑了稿件。C. K. 和 D. Z. 负责监督整个项目。

利益冲突声明

作者声明，他们不存在任何利益冲突。

数据可用性

本研究生成的所有数据集和代码均已存放于开源代码库：https://github.com/woshixuhao/Discovery_of_corona/tree/main。

参考文献

1. Kroot, M. 建立能源联系：高压输电地理学综述。《环境地理学进展》 **3**, 311–331 (2024). 2. Engel, Z. & Wszołek, T. 电晕效应导致的输电线路可听噪声：分析、建模与预测。《应用声学》 **47**, 149–163 (1996). 3. Liu, Y. 等 特高压直流输电线路电晕电流频谱特征及其与无线电干扰关系的分析。《*IEEE*电介质与电气绝缘汇刊》 **23**, 3336–3345 (2017). 4. Bian, X. 等 两种分裂导线类型的交流长期运行导线电晕产生的可听噪声和无线电噪声对比研究。《静电学杂志》 **74**, 47–55

(2015). 5. Maruvada, P. S. 高压输电线路电晕性能。（Research Studies Press，鲍多克，英国，2000）。6. Combes, M. G., Carberry, R. E., Chartier, V. L., Maruvada, P. S. & Stewart, J. R. 高压输电线路可听噪声计算方法比较。(1982). 7. Perry, D. E., Chartier, V. L. & Reiner, G. L. BPA 1100千伏输电系统开发中的电晕与电场研究。*IEEE*电力设备与系统汇刊 1728–1738 (2007). 8. Chartier, V. L. & Stearns, R. D. 架空高压交流和直流线路可听噪声预测公式。*IEEE*电力设备与系统汇刊121–130 (2007). 9. Kurniawan, I., Pischler, O. & Schichler, U. 不同架空输电线路的可听噪声计算。载于*2018*年第*53*届国际大学电力工程会议 1–6 (IEEE, 2018). 10. Wan, B., Wu, X., Pei, C., Chen, Y. & Zhou, B. 基于多变量回归法的交流输电线路声功率计算方法。《高电压工程》 **4**, 1354–1361 (2017). 11. Li, L. 等 高压直流输电线路可听噪声研究进展。《能源》（巴塞尔） **16**, 4614 (2023). 12. He, G. 等 覆冰严重程度对雾凇覆冰导线正电晕产生的可听噪声特性的影响。*IEEE*工业应用汇刊 **55**, 5269–5276 (2019). 13. Yi, Y., Chen, Z. & Wang, L. 高压直流输电线路可听噪声的统计评估与预测。载于*2016*年*IEEE*电气绝缘与介电现象会议（*CEIDP*） 279–282 (IEEE, 2016). 14. Shi C., Zhu, P. & Jiang, J. 采用五分裂导线的±800千伏特高压直流输电线路电磁环境剖面。《高电压工程》 **37**, 666–671 (2011). 15. Gilpin, L. H. 等 解释解释：机器学习可解释性概述。载于*2018*年第*5*届*IEEE*国际数据科学与高级分析会议（*DSAA*） 80–89 (IEEE, 2018). 16. Brunton, S. L. & Kutz, J. N. 机器学习在偏微分方程领域的前景方向。《自然计算科学》 **4**, 483–494 (2024). 17. Makke, N. & Chawla, S. 通过符号回归实现可解释的科学发现：综述。《人工智能评论》 **57**, 2 (2024). 18. Billard, L. & Diday, E. 符号回归分析。载于《分类、聚类与数据分析：最新进展与应用》 281–288 (施普林格, 2002). 19. Udrescu, S. M. & Tegmark, M. AI Feynman：一种用于符号回归的物理启发方法。《科学进展》 **6**, (2020). 20. Landajuela, M. 等 深度符号回归的统一框架。《进展》

神经信息处理系统s 35 (2022): 33985-33998. 21. Xu, H., Chang, H. & Zhang, D. DLGA-PDE：通过结合深度学习与遗传算法，在候选库不完整的情况下发现偏微分方程。计算物理学杂志 **418**, (2020). 22. Rudy, S. H., Brunton, S. L., Proctor, J. L. & Kutz, J. N. 偏微分方程的数据驱动发现。科学进展 **3**, (2017). 23. Sarić, A. T., Sarić, A. A., Transtrum, M. K. & Stanković, A. M. 电力系统数据驱动动态模型改进的符号回归。*IEEE*电力系统汇刊 **36**, 2390–2402 (2020). 24. Xu, H. 等 运用可解释人工智能探索地形与降水的关系，以推进未来气候预测。*Nexus*. **2**, 100045 (2025). 25. Yang, K.-H., Lee, D.-I., Hwang, G.-H., Park, J.-H. & Chartier, V. L. 利用进化计算预测架空暖通空调管线可听噪声的新公式。*IEEE*电力输送汇刊 **15**, 1243–1251 (2000). 26. Makke, N. & Chawla, S. 符号回归与可解释科学发现：综述。人工智能评论 **57**, (2024). 27. Riba, J.-R., Morosini, A. & Capelli, F. 大气空气中球-板间隙交流及正负直流可见电晕的比较研究。《能源》（巴塞尔）**11**, 2671 (2018). 28. Huang, C. 等 基于电晕放电特性预测高压直流输电线路的无线电干扰。《高电压》 **5**, 679–687 (2020). 29. Tang, J., Yang, Y., Li, yongshuang, Zhang, xiaowu & He, jinliang. 特高压交流输电线路电晕效应预测方法 I：可听噪声。《高电压工程》 **36**, (2010). 30. Cranmer, M. 使用 PySR 和 SymbolicRegression.jl 进行科学领域的可解释机器学习。*arXiv* 预印本 *arXiv:2305.01582* (2023). 31. Zhu, L. 等 激励函数在预测电晕放电引起的无线电干扰水平中的应用。等离子体科学与技术 **14**, 1091 (2012). 32. Priest, K. W., Carter, G. K. & Juette, G. W. 输电线路无线电干扰统计的计算。*IEEE*电力设备与系统汇刊92–98 (2007). 33. Moreau, M. R. & Gary, C. H. 高压输电线路无线电干扰电平的预先确定 I——激励函数的预先确定。*IEEE*电力设备与系统汇刊 284–291 (2007). 34. La Cava, W. 等 当代符号回归方法及其相对性能。神经信息处理系统进展 **2021**, 1 (2021).

补充信息：关于基于图的数据驱动发现高压输电线路电晕诱发噪声和无线电干扰的可解释规律

Hao Xu[1,2], Yuntian Chen[1,3,*], Chongqing Kang[2,*] 和 Dongxiao Zhang[1,4,*] 浙江省工业智能与数字孪生重点实验室，宁波东方理工大学，浙江宁波 315200，中国[2] 电气工程系，清华大学，北京 100084，中华人民共和国。[3] 宁波数字孪生研究院，宁波东方理工大学，浙江宁波 315200，中华人民共和国[4] 高级研究院，岭南大学，香港屯门

[*] 通讯作者。

电子邮箱：ychen@eitech.edu.cn（Y. Chen）；cqkang@tsinghua.edu.cn（C. Kang），dzhang@eitech.edu.cn（D. Zhang）。

1. 补充文本

1.1 从单相可听噪声产生到地面A计权噪声的计算

为与现场测量结果进行比较，必须将以分裂导线表面梯度和导线几何结构表示的单相可听噪声产生功率，转换为地面观测点处的A计权声压级。本小节概述本研究采用的计算流程。对于具有n根子导线、等效直径为d （厘米）且平均最大表面梯度为（E）（千伏/厘米）的分裂导线，单相单位长度的A计权可听噪声产生功率级（亦称声功率级）表示为L_{AN}，这是可听噪声预测模型进行数据驱动发现的目标变量。为获得地面某点处的A计权声压级，需在能量域中叠加所有相的贡献。对于i^{th} 相，其位置与观测点之间的水平距离和高度决定了直线距离R_i（米），相应的A计权声压级贡献为：

$$L_{p,i}=L_{AN,\mathrm{i}}-C\log_{10}R_i-5.8 \tag{S.1}$$

wh本文采用的经验模型对应的C=10 ， 以及数据驱动发现的公式对应的C=11.4

公式。$\log_{10} R_i$ 这一项表示声音从线声源经反射地面传播时的柱面扩散。假设不同相位产生的噪声以非相干叠加的方式合成，则观测点处的总A计权声压级为

$$L_p = 10 \log_{10}\left(\sum_{i=1}^{z} 1\, 0^{L_{p,i}/10}\right) \tag{S.2}$$

其中，z表示相数。对于常规三相线路，$z=3$。对于典型的三相输电塔，距离R_i 可直接根据输电塔几何结构获得。设(x_i, h_i)为第i^{th} 相导线中心相对于输电塔中心线的水平坐标和高度，(x_m, h_m)为测量点的坐标，其中$h_m \approx 1.5$ 米对应标准麦克风高度。观测点到第i相的直线距离为：

$$R_i = \sqrt{(x_m - x_i)^2 + (h_m - h_i)^2} \tag{S.3}$$

对于常见的水平排列三相塔，三根相导线大致处于相同高度h_i ，但横向偏移x_i各不相同；对于三角形排列，各相的x_i 和h_i 均不相同。无论哪种情况，一旦根据输电塔设计确定(x_i, h_i)，并给出测量位置(x_m, h_m)，即可根据式(S.3)计算距离R_i ，并代入式(S.4)。由此可得接收点位置处预测的A计权地面可听噪声级，可直接与现场测量结果进行比较。

1.2 交流输电线路中 RIEF 到无线电干扰电平的转换

无线干扰激发函数（RIEF），记为 Γ_{RI}，用于表征在规定的运行和气象条件下，分裂导线产生射频电晕电流的能力。在干扰频率 f_{RI} 下，多导体线路在频域中以单位长度串联阻抗矩阵**Z** 和并联导纳矩阵**Y**建模。这些矩阵涵盖考虑集肤效应的导线电阻和内部电感，以及根据复穿透深度获得的大地回流参数。沿线路的电压和电流相量**V**和**I**满足含分布式电流源的电报方程：

$$\begin{aligned} \frac{\mathrm{d}^2 \boldsymbol{U}}{\mathrm{d}x^2} &= \mathbf{ZYV} \\ \frac{\mathrm{d}^2 \boldsymbol{I}}{\mathrm{d}x^2} &= \mathbf{YZI} \end{aligned} \tag{S.4}$$

通过模态变换可将这些耦合方程解耦。设**M** 和**N**分别为**ZY**和**YZ**的右特征向量矩阵，其公共特征值矩阵为 λ：

$$\lambda = \mathbf{M}^{-1}\mathbf{ZYM} = \mathbf{N}^{-1}\mathbf{YZN} \tag{S.5}$$

电晕电流密度矢量为：

$$\mathbf{J}=\frac{\mathbf{C}}{2\pi\varepsilon_0}\,\Gamma \tag{S.6}$$

转换为模态量后，

$$\mathbf{J}_m=\mathbf{N}^{-1}\mathbf{J} \tag{S.7}$$

然后，考虑到

$$\mathbf{I}_m=\mathbf{g}_m\mathbf{J}_m \tag{S.8}$$

每根导线中的电晕感应电流相量可写为：

$$\mathbf{I}=\frac{\mathbf{N}\mathbf{g}\mathbf{N}^{-1}\mathbf{C}^{-1}\Gamma}{2\pi\varepsilon_0} \tag{S.9}$$

已知电晕电流后，地面观测点 $(x,0)$ 处的电磁场可通过叠加各导线产生的场求得。水平磁场分量 $H_x(x,0)$ 的计算公式为

$$H_x(x,0)=\sum_{i=1}^{N}\frac{I_i}{2\pi}\left[\frac{\mathrm{h}_i+P}{(x-x_i)^2+(\mathrm{h}_i+P)^2}\,\frac{\mathrm{h}_i-P}{(x-x_i)^2+(\mathrm{h}_i-P)^2}\right] \tag{S.10}$$

其中，I_i 为导线 i 中的复电流，(x_i,h_i) 分别为其水平位置和高度，P 为大地回流的复穿透深度。主要的垂直电场分量为：

$$E_y(x,0)=Z_0H_x(x,0) \tag{S.11}$$

其中，$Z_0\approx 120\pi\,\Omega$ 为自由空间波阻抗。对于每相包含多组导线束的情况，公式 (S.10) 明确纳入所有子导线；或者，也可以用位于分裂导线束形心处的等效导线来表示导线束，并对电流进行适当缩放。最后，观测点处的无线电干扰电平表示为高于 $1\,\mu\mathrm{V/m}$ 的电场电平，单位为分贝：

$$\mathrm{RI}(x)=20\log_{10}\left(\frac{|E_y(x,0)|}{1\,\mu\mathrm{V/m}}\right)\quad[\mathrm{dB}\,\mu\mathrm{V/m}] \tag{S.12}$$

这样一来，RIEF 函数便可在多导线线路模型中一致地传递，从而预测走廊沿线任意观测点处的无线电干扰电平。

2. 补充表格和图

表 S1. 用于预测可听噪声 A计权声压级的六种传统经验模型公式，以及通过符号回归基线发现的两种模型公式。其中，E 为平均最大导线束表面梯度（千伏/厘米），d 为子导线直径（厘米），n为导线束中的子导线数量，log表示以 10 为底的对数。这里，每个模型均以单位长度A计权可听噪声产生功率 L_{AN} 表示，其量级单位为 dB(A) re 1~μW/米。

模型名称	公式
BPA	L_{AN}=120对数 E +55 对数 d +26.4 对数 n − 128.4
ENEL	L_{AN}=85 对数 E +45 对数d +18 对数 n − 71
IREQ	升$_{AN}$=72 log 电场 +45.81 log距离 +22.71 log 导线数量 − 57.6
FGH	L_{AN}=2E+45对数 d +18对数 n − 0.3
GE	L_{AN}= $-\frac{655}{E}$+44对数 d +20 对数 n +67.9
EPRI	L_{AN}=120对数 E +54 对数 d +24.8 − 126
PySR	L_{AN}=1.58nd − 2.97n+55.6 $-\frac{915}{E}$
DSO	L_{AN}= − 2.65 $\frac{Ed}{n^2}$+62.8d $-\frac{64.8d}{\text{对数}(E)}$+0.47$n$ − 17

表 S2. 用于预测无线电干扰激发函数的五种传统经验模型公式，以及通过符号回归基线发现的两种模型公式，Γ_{RI}。其中，E 为平均最大束表面梯度（千伏/厘米），d 为子导线直径（厘米），n为导线束中的子导线数量，log表示以 10 为底的对数。

模型名称	公式
BPA	$\Gamma_{RI}=120\,\text{对数}\left(\frac{E}{15}\right)+40\,\text{对数}\left(\frac{d}{4}\right)+37.02$
CIGRE	$\Gamma_{RI}=3.5E+6d-40.69$
EPRI	$\Gamma_{RI}=\begin{cases}-\frac{580}{E}+38\,\text{对数}\left(\frac{d}{3.8}\right)+81.1, & n\leq 8\\ -\frac{580}{E}+38\log\left(\frac{d}{3.8}\right)+86.1, & n>8\end{cases}$
CISPR	$\Gamma_{RI}=70-\frac{580}{E}+35\,\text{对数}\,d-10\,\text{对数}n$
IREQ	$\Gamma_{RI}=-90.25+92.42\log E+43.03\log d-K(n)$ $K(n)=\begin{cases}0, & n=1,\\ 3.7, & n=2,\\ 6, & n\geq 3.\end{cases}$
PySR	$\Gamma_{RI}=0.51(d+158.407)\ln(\ln((n-6.35)E))-613$
DSO	$\Gamma_{RI}=11.1\frac{E}{n}+18.2d-68.6\frac{d}{n}+0.99n-16.1$