日前电价预测的混合模型：结合基本面建模与随机建模

Mira Watermeyer[a], *, Thomas Möbius[b], **, Oliver Grothe[a] 和 Felix Müsgens[b]

[a]卡尔斯鲁厄理工学院（*KIT*），运筹学研究所，卡尔斯鲁厄，*76131*，德国

[b]勃兰登堡工业大学（*B-TU*），能源经济学教席，科特布斯，*03046*，德国

arXiv:2304.09336v1 [econ.EM] 18 Apr 2023

文献信息

关键词：电价预测 混合模型 能源系统建模 随机建模 误差改进 概率预测

摘 要

短期电价的准确预测对于制定有效的交易策略、电厂调度、利润最大化以及高效的系统运行至关重要。然而，供需方面的不确定性使此类预测面临巨大挑战。为应对这一挑战，我们提出了一种混合模型，将技术经济能源系统模型与随机模型相结合。在我们的混合方法中，技术经济模型能够深入理解市场。它能够捕捉影响电价的潜在因素及其作用，而仅依靠统计模型则无法实现这一点。统计模型则通过对价格的解读，将非技术经济因素纳入考量，例如市场参与者的预期和投机行为。该混合模型既能生成传统的点预测，也能生成概率预测，从而全面呈现市场格局。概率预测尤为重要，因为它能够考虑市场不确定性，有助于制定明智的决策和实施风险管理。我们的模型取得了最先进的成果，帮助市场参与者做出明智的决策，并更高效地运行其系统。

1. 引言

能源行业的准确预测对于包括行业从业者、研究人员和政策制定者在内的众多利益相关者至关重要。金融和运营决策以及监管干预的有效性取决于对相关领域未来发展的准确预测。因此，电价预测已成为重点关注领域 [Weron, 2014]。随着电力行业因放松管制和自由化而面临日益激烈的竞争，日前电价预测以及对次日市场状况的洞察，对于制定能够最大化企业利润率并确保可靠电网运行的竞价策略和生产计划至关重要。近年来，由于可再生能源不断发展，以及需要将其纳入能源系统，同时基础设施挑战增加、大宗商品价格波动，能源市场中的不确定性日益加剧，因此量化不确定性变得越来越重要 [Nowotarski 和 Weron, 2018年, Hong 等, 2016, 2020]。概率预测有助于能源系统的规划和运行，使人们能够评估不确定性，并在各种可能的未来事件背景下制定未来策略 [Amjady 和 Hemmati, 2006]。

本文提出了一种新颖的开源混合模型，通过结合两种主要的方法流派——技术经济能源系统建模和随机建模，对日前电价进行点预测和概率预测。技术经济模型是基础性的能源系统模型，通过对能源系统进行自下而上优化来确定（部分）市场均衡。它们能够通过识别能源市场中的技术经济相互依赖关系，解释实际发展情况并反映结构突变。然而，在估计

mira.watermeyer@kit.edu（M. Watermeyer）；thomas.moebius@b-tu.de（T. Möbius）

https://www.as.ior.kit.edu/（M. Watermeyer）；https://www.b-tu.de/fg-energiewirtschaft/（T. Möbius）

ORCID:

短期内（例如日前、日内）的电价时，与其他模型类别相比，这些模型会产生更大且更具系统性的误差。另一方面，随机模型从历史中学习，并利用历史数据进行开发和训练，从而能够捕捉市场中的波动和不确定性，尤其是在短期内。它们具有很高的灵活性，并能够指定预测范围和分布。不过，由于依赖历史数据，它们只能在事后捕捉结构突变以及外部影响因素的变化。

我们提出的混合模型结合了技术经济能源系统模型和随机模型的优势，旨在开发一种更稳健、更准确的日前市场电价预测方法。该模型保留了技术经济能源系统模型的结构性表述，从而保留了对市场驱动机制的洞察，同时融入随机短期结构和分布函数，以考虑不确定性。该模型采用最先进的方法生成点预测和概率价格预测。这些概率电价预测为每个潜在价格情景提供相应概率，因此对行业的价值日益提升（例如，在评估负价格出现的概率或评估价格预测的总体风险水平时）。

该模型的示意图见图1。模型采用滚动窗口方法。在每次迭代中，模型仅使用日前市场收市前已知的数据来预测日前价格，从而准确反映在这些市场中作出决策的利益相关者所掌握的信息。模型每天重复应用（d_n），以生成日前市场未来24小时的预测。每次日度预测包括四个步骤，即随机数据预处理、参数密度预测、能源系统优化和随机数据后处理，以生成点预测和概率预测。

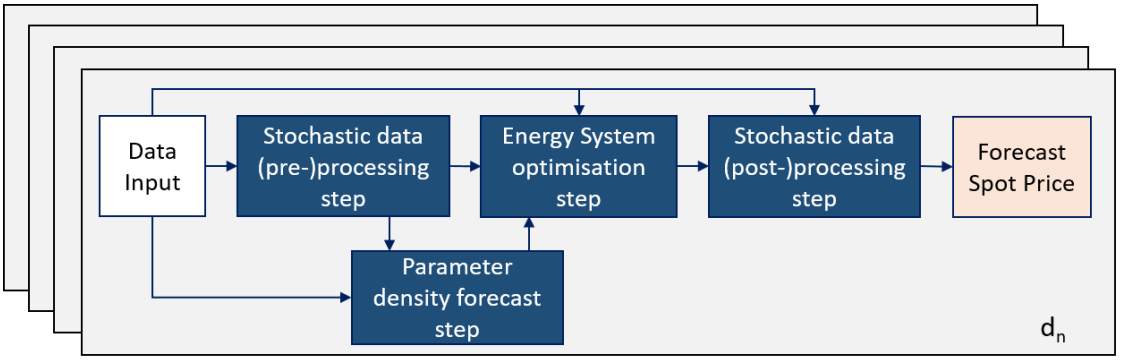


图 1： 混合模型示意图。

第一步，即随机数据预处理，旨在能源系统优化步骤之前提高输入数据的准确性，并为参数密度预测步骤奠定基础。在第二步中，参数密度预测为第三步能源系统优化所选的输入参数生成预测区间，使我们能够在市场参与者的运营决策中考虑不确定性。该步骤会采用第一步和第二步改进后的输入数据。它构建一个随机优化模型，使系统总成本最小化，确定供需平衡，并计算小时边际系统成本，这些成本可解释为价格估计器。

这些价格估计器是第四步也是最后一步——随机数据后处理的初始值。模型利用多维模型对价格估计器的误差进行映射，并通过组合单变量和多变量方法捕捉季节性影响及其结构，从而得到增强的价格预测。通过对价格预测误差进行建模和改进，该模型计算预测价格的预测区间和概率密度，从而实现不确定性量化。最后，我们的模型能够结合两类方法的优势，获得卓越的最先进的价格预测，包括点预测和概率预测，以捕捉市场的随机不确定性。

本文从三个主要方面为相关研究作出了贡献。首先，本文提出了一种新颖的混合模型，为结合技术经济模型与随机能源模型提供了通用框架。由于该模型的源代码和代数表达式已完整公开于网上，其他研究者可以运用我们的方法论，并将其拓展到世界各地不同的时间段和电力市场。其次，本文证明，技术经济能源系统模型能够有所贡献。

对于短期价格预测而言，尤其是在与随机模型结合以实现误差改进时。我们的混合模型在技术经济模型所提供的洞见基础上，能够生成高度准确的日前电价预测。我们通过一项基于欧洲数据、聚焦于德国这一欧洲最大的电力市场的实证分析，展示了该模型的价值。第三，我们的混合模型除了提供点预测外，还能提供概率预测，使电厂运营商能够例如量化任意给定小时价格变为负值的可能性。

本章其余部分安排如下。第 2 节回顾现有文献。随后，我们在第 3 节提供使用模型和复现本研究所需的全部信息。第 4 节介绍我们的方法论。接下来，第 5 节展示并评估混合模型的结果，而第 6 节则深入分析模型单个步骤。最后，第7节给出一些总结性意见。

2. 文献

由于要在预测中实现高准确性所面临的复杂性和非凡挑战，关于电价的研究引起了众多学者的兴趣。他们开发并完善了大量方法，以实现高准确性并适应电力市场的变化。过去二十年间，相关出版物的数量迅速增加。

Weron [2014] 详细回顾了电价预测的多种方法，包括以下五类模型：多智能体模型、基本面模型、简约模型、统计模型和人工智能模型。此前，Aggarwal 等 [2009] 概述了电价预测所采用的方法。不过，他们的重点是随机时间序列、因果模型和基于人工智能的模型。Weron 和 Ziel [2019] 以及Hong 等 [2020] 对能源预测进行了总体回顾并展望了未来。关于预测理论与实践的最新综述来自Petropoulos 等 [2022], , 他们概述了用于制定、组织和评估预测的各类理论模型、方法、原则和途径。此外，他们还提供了若干现实案例，展示这些理论概念的应用方式。

许多出版物基于时间序列模型开展时间序列分析，这类模型尤其适用于短期电价预测。时间序列模型是回归模型的一种特殊子类型，其中目标变量 y 由时间序列的过去值等作为回归变量来表示 x。这类模型包括自回归移动平均模型（ARMA）、广义自回归条件异方差模型（GARCH）和马尔可夫状态转换模型（MS）。Steinert 和 Ziel [2019], 例如，构建了一个包含 24 个独立模型的自回归模型，即每天每个小时对应一个模型；该模型还纳入了电力期货价格，以生成逐小时的电价预测。Nowotarski 和 Weron [2016] 提到，将数值分解为不同成分是时间序列分析中的常见做法，并表明，将一组电价分解为长期季节性成分和随机分量，对二者分别建模并组合其预测结果，能够显著提升时间序列模型的质量。在一项广泛的研究中，Ziel and Weron [2018] 比较了使用高频数据集进行时间序列建模时的两种方案。他们比较了单变量模型框架和多变量模型框架：在单变量模型框架中，为整个时间序列构建一个模型；而在多变量模型框架中，每天的每个小时都分别独立建模。这种做法源于电力市场采用日前拍卖的组织方式，例如美国或欧洲的电力市场。研究显示，没有任何一种框架明显占优，这表明组合两种建模方法可能有助于提高预测准确性。

虽然Steinert 和 Ziel [2019], Nowotarski 和 Weron [2016] 以及Ziel and Weron [2018] 总体上关注日前电价，Christensen 等 [2012] 使用自回归条件风险模型的非线性变体来预测价格峰值，并将其作为一个离散时间点过程进行整体处理，Eichler 等 [2014]采用基于自回归条件风险率模型的方法，而Manner 等 [2016]则在动态多元二元选择模型中刻画不同电力市场之间的区域间联系，以预测电价尖峰。Garcia 等 [2005] 建立了用于预测日前电价的GARCH模型，而Hickey 等 [2012]评估了ARMAX-GARCH模型对美国短期价格的预测准确性，确定模型选择在很大程度上取决于地点、预测期限和监管；其中，非对称幂自回归条件异方差（APARCH）模型更适用于放松管制市场，而GARCH模型则更适用于

受监管市场。Bordignon 等 [2013] 建立了线性回归模型，以解释价格与各种价格驱动因素之间的关系，并使用时变参数（TVR）模型和MS模型来捕捉峰值和不连续性。应用MS模型的其他例子包括Kosater 和 Mosler [2006]针对德国市场的研究，以及Bierbrauer 等 [2004] 针对北欧市场的研究。值得注意的是，在一篇近期论文中，Mari 和 Mari [2022] 使用基于深度学习的状态转换模型来预测电价。

参数丰富的ARX模型代表一种特殊类型的时间序列模型。由Uniejewski 等人 [2016] 引入的套索估计自回归（LEAR）模型由Lago等人 [2021]进一步发展。
为评估电价预测领域未来的模型开发，并比较最先进的统计方法和深度学习方法，提供一套最佳实践，Lago等人 [2021] 根据最新研究成果定义了深度神经网络（DNN）和 LEAR 模型。结合各种评估指标，这些模型可通过 Python工具箱获取，用于评估新算法。因此，我们将混合模型与统计基准模型进行比较。

深度学习模型（例如人工神经网络（ANN）、DNN、长短期记忆（LSTM）网络、循环神经网络（RNN）、前馈神经网络）在电价预测中的应用占比不断提高。除了所引基准模型外，Panapakidis 和 Dagoumas [2016], 例如，研究了采用不同输入和人工神经网络类型的 ANN。这些作者认为，此类模型功能全面且灵活性高。在分析不同市场之间相互影响时，Lago等人 [2018] 开发了一种 DNN，考虑互联市场的特征以提高预测准确率。值得注意的是，他们表明，同时预测两个市场的价格可以提高预测准确率。 Amjady [2006] 开发了一种模糊神经网络，用于预测西班牙日前市场的小时电价。值得注意的是，深度学习与时间序列模型的结合体现在 Marcjasz 等人 [2019]提出的非线性自回归神经网络中。

为比较使用外生回归变量的时间序列模型和神经网络模型，Lehna等人 [2022] 采用四种不同的方法预测德国日前电力市场：季节性自回归积分滑动平均（(S)ARIMA(X)）模型、LSTM神经网络、卷积神经网络长短期记忆模型（CNN-LSTM）以及扩展的两阶段多变量向量自回归（VAR）模型。LSTM模型的平均准确性最高，而两阶段VAR模型在较短的预测时域上更具优势。两种方法组合后，其准确性优于各自单独使用时的表现。
迄今为止介绍的方法，基本上都基于历史日前电价时间序列。另一种方法是使用模拟个体市场参与者（代理）行为的模型。Qussous等人 [2022],例如，他们开发了一种基于代理的模型，用于推导日前价格并模拟市场参与者的竞价策略。为评估该模型，他们再现了2016至2019年德国竞价区的日前电价。与其他用于短期电价预测的技术经济方法相比，该基于代理的模型准确性最高。因此，我们将本文提出的混合模型与这一模型进行比较。

由于技术经济能源系统模型能够解释有效市场结果，并对整个电力系统进行全面建模，因此已被用于电价的事后分析。 Muesgens [2006等人] 和 Borenstein等人 [2002] 分别复现日前价格，以评估德国和美国是否存在市场势力。 Keles等人 [2013] 研究了适当的风电馈入时间序列对于在电价模拟中获得更好结果的重要性。 Hirth [2013]确定了可再生能源的市场价值，而 Sensfuß等人 [2008] 通过计算日前价格量化了优先次序效应。优先次序效应描述了由于可再生能源的边际成本较低，可再生能源取代化石燃料发电，以及随后总电力成本下降的现象。 Pape等人 [2016] 分析了技术经济能源系统模型在多大程度上能够解释和表示日前和日内电价。然而，值得注意的是，与其他方法相比，他们表明该方法在解释短期电价方面存在显著缺陷。

与对已经发生的实际价格进行考察的事后电价分析不同，本文聚焦于事前预测。用于事前预测的技术经济能源系统模型有一项关键缺点：它们不会利用近期历史价格来校准价格估计器。因此，与“从过去学习”的计量经济模型相比，它们可能难以解释随机的短期波动。然而，能源系统模型具有一项优势，即它们以成熟的经济理论为基础，并能重现

市场的运行方式。因此，它们能够独立于过去的数据预测价格，也较少受到市场结构变化及其他类似因素的影响。与此相应，我们没有发现技术经济能源系统模型被用于短期电价事前预测的案例。不过，这类模型已用于长期电价的事前预测，因为随机的短期波动在长期电价中不那么显著 [例如，Muesgens，2020年；Green 和 Vasilakos，2010年；Lamont， 2008]。此外，技术经济能源系统模型也应用于价格估计以外的文献领域（例如，确定需求响应的价值 [Misconel 等，2021年；Kirchem 等， 2020], ；确定最优输电扩建计划 [van der Weijde 和 Hobbs， 2012], ；支持市级决策 [Scheller 和 Bruckner， 2019], ；分析电转气的影响 [Lynch 等， 2019], ；评估减少 CO_2 排放的政策工具 [Sgarciu 等，2023]）。此外，Plaga 和 Bertsch [2023] 深入考察了能源系统模型如何纳入气候不确定性。Ventosa 等人 [2005]对能源系统建模进行了全面综述。

近年来，混合方法在电价预测领域受到了广泛关注。混合模型是指结合两种或更多种不同方法的模型。其旨在利用所采用方法的综合优势，同时减轻各自的不足，从而取得更好的整体结果。人们已经开发出许多结合多种方法的混合方法。Aggarwal和Tripathi [2017], 例如，提出了一种混合方法，利用小波变换、时间序列时延神经网络和误差预测算法，预测ISO New England市场的日前电价。Chang等人 [2019] 将Adam优化的LSTM神经网络与小波变换相结合，前者用于生成电价，后者用于将电价时间序列分解为若干电价序列。Cheng等人 [2019]提出了经验小波变换、支持向量回归、双向长短期记忆网络和贝叶斯优化的组合。Nazar等人 [2018]将三阶段混合模型应用于Nordpool的DK2区域和西班牙电力市场。第一阶段采用小波变换和卡尔曼机，将价格数据分解为不同的频率成分。第二阶段使用自适应神经模糊推理系统（ANFIS）预测价格频率成分。在第三阶段，将第二阶段的输出输入ANFIS，以提高预测准确率。Yang等人 [2017],将小波变换和ARMA与基于核的极限学习机相结合，Olamaee等人 [2016]则将二者与径向基函数神经网络相结合。Zhang等人 [2020] 提出了一种基于变分模态分解、自适应粒子群优化、季节性自回归积分滑动平均模型和深度置信网络的混合模型，用于短期电价预测。

上述大多数混合模型采用统计方法和深度学习方法。然而，少数应用还将技术经济能源系统模型与另一种方法相结合。例如，de Marcos 等 [2019]详细介绍了一种面向伊比利亚市场的短期混合电价预测模型，该模型将技术经济成本发电优化模型与 ANN 相结合。Gonzalez 等 [2012] 提出了两种基于技术经济电力市场模型的混合方法。他们聚焦于英国日前市场，分别将该模型与另外两种模型相结合：一种是含价格驱动因素外生数据的线性自回归模型（ARX模型），另一种是带平滑转换的非线性逻辑回归模型（LSTR模型）；后者在结构变化期间体现为机制转换。他们的结果支持这样一种观点：纳入基本面信息有助于提高价格预测准确度。尤其是在高度波动的时期，非线性混合模型能够取得更好的结果。在Möbius 等 [2023], 我们此前的研究中，我们引入了一种针对日前市场定制的技术经济市场模型，并将其与随机模型相结合，以提高日前负荷预测在日前电价估计中的准确性。我们强调了更准确的负荷预测对能源系统模型生成的日前价格估计器所产生的积极影响。然而，这种方法仅代表第一步；它尚未充分发挥混合模型的潜力，即无缝整合技术经济模型与随机模型各自的优势。

关于电价预测的文献主要侧重于开发日前市场的点预测方法。然而，近年来，人们对概率预测方法的兴趣日益浓厚 [Hong 等， 2020]全球能源预测竞赛（GEFCom2014） [参见Hong 等， 2016] 推动了这一趋势，此后，已有许多关于这一主题的研究发表。Nowotarski 和 Weron [2018]全面概述了该领域采用的不同方法。Maciejowska 和 Nowotarski [2016] 针对 GEFCom2014 开发了一种分四个步骤结合点预测与概率预测的混合模型。

概率电价预测的常见方法包括使用时间序列模型，例如自回归积分滑动平均模型、广义自回归条件异方差模型和指数平滑模型（ETS） [例如，Weron 和 Misiorek， 2008]以及使用深度学习模型。自助法广泛用于结合深度学习方法 [例如，Chen 等人，2012年，Wan 等人，2014年，Rafiei 等人，2017年，Khosravi 等人， 2013]。在深度学习的基础上，Zhao 等人 [2008] 使用支持向量

机（SVM）估计预测区间和密度预测，而Zhou 等人 [2006] 使用扩展的 ARIMA 模型进行相同的工作。Panagiotelis 和 Smith[2008]提出了一种用于概率预测的计量经济模型。Manner 等人 [2019] 使用藤 copula 模型预测来自互联电力市场的日前电价向量的分位数，而Grothe 等人 [2023]提出了一种基于 copula 技术的方法，通过对跨小时依赖关系进行建模来生成多元概率预测。与点预测相比，在概率预测中考虑这些依赖关系并不常见。然而，将这些依赖关系纳入生成概率电价预测的方法论中，可以提高预测准确度。

历史模拟和基于分布的预测区间是其他常用方法。历史模拟利用历史数据模拟多个情景，以此估计风险并生成预测区间；随后利用结果估计不同结果发生的概率 [例如，Weron 和 Misiorek，2008年，Nowotarski 和 Weron，2015]。基于分布的预测区间则根据历史数据的分布计算 [例如，Misiorek 等人，2006，Zhao 等人，2008年，Dudek，2016年，Maciejowska 等，2016年，Panagiotelis 和 Smith， 2008]。Weron 提供了关于基于分布和历史模拟生成预测区间的理论介绍Weron [2006]。

分位数回归平均（QRA）是近年来在概率电价预测领域日益受到关注的一种方法。它结合多个分位数回归模型的预测，每个模型都经过训练，用于预测响应变量的不同分位数。该方法最初由 Nowotarski 和 Weron[2015] 正式提出，此后因其准确性而持续得到应用和进一步发展 [例如，Maciejowska 等，2016年，Nowotarski 和 Weron，2014年，Marcjasz 等人，2020年，Uniejewski 等人，2019年，Uniejewski 和 Weron， 2021]。

尽管概率预测在各种模型中的重要性日益凸显，但目前仍普遍缺乏将概率预测与技术经济能源系统模型相结合的方法。本文旨在填补这一研究空白。通过针对短期电力市场调整和开发能源系统模型，并将该模型通过多个步骤与常见的随机模型相结合，我们能够发挥两类模型各自的优势，并为能源系统模型开展短期电价预测开辟新的领域。在已经突出展示将随机模型（用于获得更准确的负荷预测）与能源系统模型（用于日前市场，由 Möbius 等 [2023]开发）相结合的积极作用之后，这些构建模块被纳入混合模型中。我们证明，多层混合模型能够利用技术经济模型和随机模型进行点预测和概率电价预测。

3. 数据

在本研究中，我们开发了一种混合模型，将随机建模方法与能源系统优化相结合，以预测批发电价。值得注意的是，这种能源系统优化需要多个输入。表1概述了所需的输入数据。本节将进一步介绍数据的获取方式及其应用方法。

尽管我们的建模方法可以应用于许多市场，但我们的实证分析仅聚焦于德国现货市场。然而，欧洲电力市场之间的高度一体化及由此产生的相互依赖关系，要求对这些市场进行全面表征，尤其是在能源系统优化步骤中。图2 展示了所收集数据的地理范围以及欧洲市场之间的互联情况。我们考虑了欧盟27个成员国中大多数国家的竞价区[1] ，以及挪威、瑞士和英国。[2] 除非另有说明，所收集的数据涵盖2016年至2020年。

电力需求由系统每小时的电力负荷值表示，其中日前预测值和实际值均由相应的输电系统运营商（TSO）发布，并由ENTSO-E Transparency Platform [2021a]提供。德国-卢森堡竞价区收集的负荷数据涵盖2015年–2020年。在能源系统模型中，电力需求通常被认为在短期内具有波动性且缺乏弹性

[1]保加利亚、塞浦路斯、希腊、冰岛、爱尔兰、马耳他和罗马尼亚未纳入其中。[2]需要注意的是，我们将西班牙和葡萄牙的竞价区汇总为单一的“伊比利亚”市场，并将立陶宛、爱沙尼亚和拉脱维亚的竞价区汇总为单一的“波罗的海”市场。此外，需要注意的是，我们保留了各国境内不同的竞价区。然而，我们对以下区域进行了汇总：挪威的NO1–NO5区域、瑞典的SE1–SE3区域，以及意大利除IT-North以外的所有区域。

表 1所需数据概览

参数	来源
CO_2 价格	Sandbag [2020]
调频电力采购	Regelleistung.net [2018]
发电容量的效率	Schröder等人 [2013], 开放电力系统数据 [2020a]
部分负荷下的效率损失	Schröder等人 [2013]
电力需求（负荷）	ENTSO-E 透明度平台 [2021a]
储能设施的能量功率比	自行假设：9
燃料价格 （褐煤、核能、煤炭、天然气、石油）	德国联邦统计局 [2021], ENTSO-E [2018], ENTSO-E [2018], EEX [2021]
发电与储能容量	德国联邦网络局 [2021],德国联邦环境局 [2020], EBC [2021], ENTSO-E 透明度平台 [2021b], 开放电力系统数据 [2020b]
热电联产机组的发电量	欧盟委员会 [2021]
历史发电量	ENTSO-E 透明度平台 [2021c]
负荷削减成本	自行假设：3,000 欧分/兆瓦时
最低出力水平	Schröder等人 [2013]
NTCs	ENTSO-E 透明度平台 [2021d], JAO联合分配办公室 [2021]
可变运行维护成本	Schröder等人 [2013]
发电厂停运	ENTSO-E 透明度平台 [2021e]
可再生能源馈入	ENTSO-E 透明度平台 [2021f]
可再生能源弃电成本	自行假设：20 欧分/兆瓦时
启动成本	Schröder等人 [2013]
水电季节性可用性	ENTSO-E 透明度平台 [2021c]
温度（日均）	开放电力系统数据 [2020a]
水价值	ENTSO-E 透明度平台 [2021c], ENTSO-E 透明度平台 [2021g]

这一术语。然而，在供应短缺时，通常可以选择切负荷。在我们的应用中，我们假设切负荷成本为 3,000 =欧分/兆瓦时，因为这是 2022[3]年9月之前的最大报价上限。

间歇性可再生能源，即陆上风电、海上风电和光伏发电（PV）的可用性取决于气象条件，并且每小时都会发生变化。这些能源来源的馈入数据由ENTSO-E 透明度平台 [2021f]以日前预测的形式按小时提供。尽管依赖天气，可再生能源发电仍可能被有意弃用。鉴于德国和欧洲针对可再生能源发电的各种支持机制会阻止可再生能源来源在出现负价格时立即停运，本研究假设弃电成本为 20 =C/MWh_{el}。

对于传统火力发电，我们区分了十种技术，并在其技术参数（尤其是影响效率的参数）随时间发生显著变化时，进一步按投运年限细分。我们采用 30 个容量集群，根据技术类型和投运日期对发电厂进行分组。我们从 开放电力系统数据 [2020b] 数据中推导出基于技术类型和投运年限的效率，并将其分配给相应的容量集群。

此外，我们根据Schröder等人 [2013]的研究，为各集群设定最低出力水平以及部分负荷运行时的效率损失。因此，当供给跟随需求和可再生能源的波动时，由于物理约束（最低出力水平）和经济激励（效率损失），机组会被激励停机。德国市场机组的容量、燃料类型、发电技术和投运日期数据取自德国联邦网络局 [2021], 德国联邦环境局 [2020] 和开放电力系统数据 [2020b]。对于研究中考虑的其余市场，我们使用来自 ENTSO-E 透明度平台 [2021b], 开放电力系统数据 [2020b] 和EBC [2021]的数据。

[3]2022 年 9 月 20 日，最大报价上限提高至 4,000 =欧分/兆瓦时。

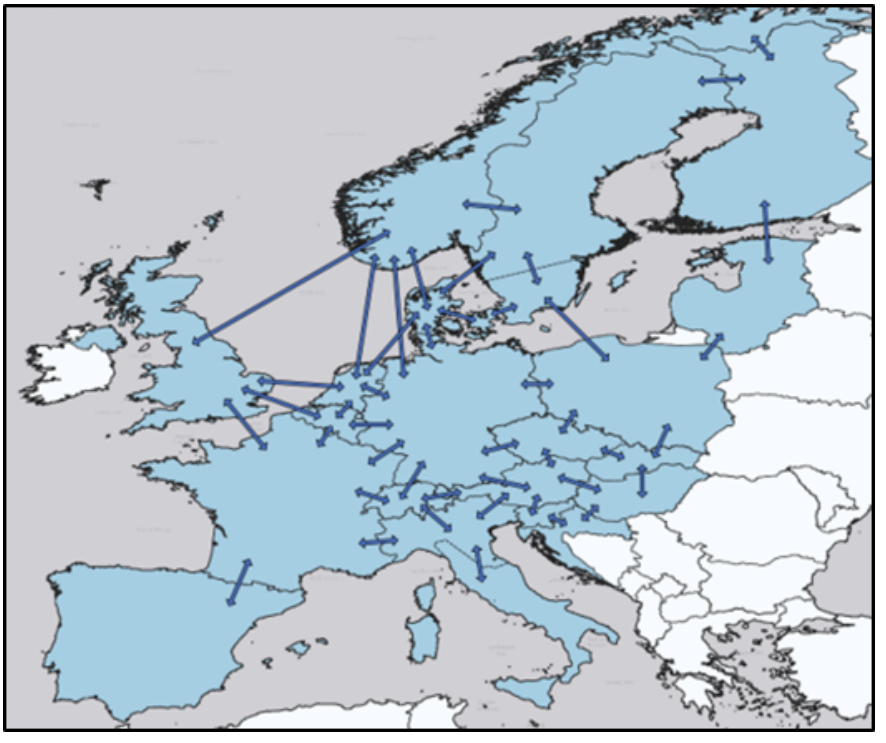


图2：能源系统模型的地理范围。

电厂效率以及燃料和 CO_2 排放成本构成了常规火电技术的可变发电成本。燃料成本方面，我们采用由以下机构提供的日度天然气价格：EEX [2021],、取自以下机构的月度煤炭价格：德国联邦统计局 [2021] 以及来自以下机构的月度石油价格：德国联邦统计局[2021]。核能和褐煤的燃料成本取自 Entso-e [2018]，并假设在本研究的预测期限内保持不变。CO_2 排放配额价格采用以下机构提供的周度数据：Sandbag [2020]。

由于发电厂在启动过程中需要耗费时间并使用额外燃料进行预热，燃料和 CO_2 价格也会影响发电厂启动的成本。有关启动的进一步数据（例如辅助燃料消耗、折旧）来源于Schröder等人 [2013]。

发电量取决于发电厂的装机容量和技术可用率。因此，我们考虑日前市场关闭前已知的所有计划停运和非计划停运。每小时停运信息获取自ENTSO-E 透明度平台 [2021e]。

在大多数电力市场中，热电联产（CHP）电厂得到应用，其中发电与供热相互关联并相互依赖。为反映这种关系，我们对所有热电联产机组施加必运行条件，以确保其在由热需求确定的最低出力水平下运行。这些出力水平分两步确定。首先，我们计算每小时的热需求系数，该系数由温度相关（空间供暖）和温度无关（生活热水和工艺热）两部分组成。温度驱动的热需求采用基于从开放电力系统数据获取的平均温度数据推导出的采暖度日计算开放电力系统数据 [2020a], ，而温度无关的热需求则根据由 Hellwig [2003]提供的每小时和每日用能模式获得欧盟委员会 [2021]。

调节电力对于系统运营商确保始终维持频率稳定性至关重要。日前市场受到调节电力供应市场的影响，这意味着预留调节电力容量无法投放到日前市场。需要采购的调节电力数量是取自Regelleistung.net [2018]的招标结果均值。

除常规火电技术和间歇性可再生能源外，我们还考虑废弃物、生物质、储能、水库水电和径流式水电。由于废弃物和生物质的运行在历史上基本保持稳定（参见ENTSO-E 透明度平台 [2021c]），我们将这两种技术均作为基荷处理。

储能单元分为高容量能量比和低容量能量比两类。高容量能量比的储能单元会主动充电和放电。我们仅考虑该组中的一部分抽水蓄能电站（PSP）。这些抽水蓄能电站的涡轮机总容量详见ENTSO-E透明度平台 [2021b]，，储能循环的效率约为75 % [Schröder等人， 2013]。对于这些抽水蓄能电站，我们假设容量能量比为1/9 [参见 德国能源署， 2010]。这意味着，在储能耗尽之前，电站可以满负荷发电九小时。低容量能量比的储能单元包括长期抽水蓄能电站以及水库水电。在模型中，我们为其分配可变发电成本（即用水量的机会成本）。利用ENTSO-E 透明度平台 [2021g] 中的历史电价，以及ENTSO-E 透明度平台 [2021c],中相应时段观测到的发电和抽水活动，我们构建了长期抽水蓄能电站和水库水电的阶梯式merit order。径流式水电和高容量能量比的抽水蓄能电站则根据ENTSO-E 透明度平台 [2021c]的历史发电数据，分配月度可用率。我们假设，70 %的抽水蓄能容量属于高容量能量比抽水蓄能电站，其余30 %的抽水蓄能容量属于低容量能量比抽水蓄能电站。

欧洲电力市场通过跨境输电容量实现了高度互联。因此，德国电力市场也融入了欧洲电力系统，互联线总容量为27吉瓦，超过德国峰值负荷的30 %。[4] 年度出口总量约占德国2019年年用电量的13%，进口总量约占同年用电量的7%，二者均规模可观。在我们的能源系统模型中，净传输容量（NTC）限制了市场区域之间的输电。我们采用由ENTSO-E 透明度平台[2021d] 和JAO联合分配办公室 [2021]提供的逐小时日前净传输容量预测。

尽管参数化数据集受到众多利益相关者的关注，但复现这些数据集需要付出巨大努力 [例如，参见Schröder等人，2013年，Kunz等人， 2017]；因此，我们的输入数据可在补充材料中找到。

4. 方法论

整个混合模型的实现均为开源；我们能够公开的所有数据均可通过以下链接在GitHub上获取：
https://github.com/ProKoMoProject/A-hybrid-model-for-day-ahead-electricity-price-forecasting-combining-fundamental-and-stochastic-mod。

<escape>本节介绍混合模型的各个组成部分，首先介绍第 4.1 节和第 4.2 节中数据预处理和参数密度预测步骤所采用的方法。接下来，在第 4.3 节中，介绍用于在能源系统优化步骤中生成首批价格估计器的调度市场模型。最后，在第 4.4 节中，详细介绍随机后处理步骤所采用的模型。</escape>

4.1. 随机数据预处理步骤

使用调度模型对电力市场进行建模，需要了解多个不同的基本变量，其中包括需求（以负荷表示），这是一个关键要素。输电系统运营商提供实际负荷和日前负荷预测，正如例如Maciejowska 等 [2021] 和 Möbius 等 [2023]所展示的，这一预测仍有改进空间。为改进该预测，我们使用随机模型进行数据预处理。此外，我们还对提前两日负荷预测进行建模，以根据第二个后续日的负荷捕捉发电厂启动和停机，并以输电系统运营商的日前负荷预测作为起点。

日前负荷预测

我们采用 Möbius 等 [2023] 最初提出的方法来改进负荷预测。因此，我们有时会引用该方法，以提供详细的规范和分析。我们提出了一种纯内生的时间序列方法：一个用于

[4]请注意，互联线路的可用性取决于多种因素（例如市场区域内的拥堵情况）。

输电系统运营商的负荷预测误差 ε_t 仅取决于预测误差本身的过去值，进而取决于输电系统运营商的负荷预测 $\hat{l}_t$。预测误差是实际负荷数据 l_t 与输电系统运营商的负荷预测 $\hat{l}_t$之差。对误差建立模型并进行预测，可以改进输电系统运营商的负荷预测值。时刻 t 的负荷预测值 $\hat{l}_t^*$ 由下式给出：

$$\hat{l}_t^* = \hat{l}_t + \hat{\varepsilon}_t, \tag{1}$$

其中， $\hat{\varepsilon}_t$ 是我们预测的输电系统运营商负荷预测误差。因此， $\hat{l}^*$ 是改进后的负荷预测值，其中我们根据其误差的可预测结构对原始预测值进行了调整。子索引 t 表示连续小时。

为对预测误差进行建模和预测 ε_t，我们使用分解模型，将误差时间序列分解为季节性分量和随机分量之和（有关时间序列模型的全面介绍，请参见Lütkepohl [2005],Hyndman 和 Athanasopoulos[2021], Box 等人 [2015]）：

$$\varepsilon_t = SC_t + RC_t, \tag{2}$$

其中， SC_t 是时间 t处的季节性分量， RC_t 是剩余的随机分量。

通过捕捉周季节性，时间 t 的季节性分量 SC_t 根据公式（3）定义，其中 $HS^{h,\ wd}$ 是小时 $h = 1$至 24 以及日期 $wd = 1$（星期一）至 7（星期日）对应时段的输电系统运营商预测误差的平均值，而 $HoW^{h,\ wd}_{t}$ 表示用于处理日内小时和星期几的虚拟变量：

$$SC_t = \sum_{h=1}^{24} \sum_{wd=1}^{7} HoW_t^{h,wd} \cdot HS^{h,wd}, \tag{3}$$

with

$$HS^{h,wd} := \frac{\sum_{s=t-h-24}^{t-h-l_w-23} \varepsilon_s \cdot HoW_s^{h,wd}}{\sum_{s=t-h-24}^{t-h-l_w-23} HoW_s^{h,wd}}, \tag{4}$$

$$HoW^{h,\ wd}_{t} = \begin{cases} 1, & \text{如果 } t \text{ 是每周第 wd 天的第 h 小时，否则。} \\ 0, & \end{cases}$$

对于时间序列的残差分量 $RC_t = \varepsilon_t - SC_t$，我们提出一个计量经济学 SARMA $(1,1)\text{x}(1,1)_{24}$模型，其表达式如下：

$$\begin{aligned} RC_t =& \phi_0 + \phi_1 \cdot RC_{t-1} + \phi_{24} \cdot RC_{t-24} - \phi_1\phi_{24} \cdot RC_{t-25} \\ &+ \omega_1 \cdot \psi_{t-1} + \omega_{24} \cdot \psi_{t-24} + \omega_1\omega_{24} \cdot \psi_{t-25} \\ &+ \psi_t, \end{aligned} \tag{5}$$

其中假设创新项具有同方差性且服从正态分布。该模型包含一个额外的 24 小时季节性分量，因此具有随机性、灵活性，并取决于过去小时和日期的数值。相比之下， SC_t 描述的是静态季节性。

模型在为期一年的校准窗口内进行估计。在混合模型中，校准窗口按整天不断滚动，并递归地预测次日的 24 小时。如果由于时间点位于未来而没有可用的实际数据，我们会使用模型变量的预测值。

提前两日负荷预测

发电厂会根据当天的需求、前一天的需求以及预计的次日需求，作出启动和停机决策。为在调度模型中考虑这一点，需要提前两天预测负荷消耗。我们使用当前负荷的建模与预测 l_t （即

$\hat{}$输电系统运营商的日前负荷预测 l_t）作为起点，预测第二个、后天的日前预测，从而得到提前两天的预测$\hat{l}_t^{2DA} = \hat{l}_t$。

对于该模型，我们提出一种计量经济学 SARMA $(1,1)\mathrm{x}(1,1)_{24}$ 模型，并加入一个额外的外生变量：滞后项为 168 的输电系统运营商负荷预测：

$$\begin{aligned}\hat{\hat{l}}_t =& \phi_0 + \phi_1 \cdot \hat{l}_{t-1} + \phi_{168} \cdot \hat{l}_{t-168} + \phi_{24} \cdot \hat{l}_{t-24} - \phi_1\phi_{24} \cdot \hat{l}_{t-25} \\ &+ \omega_1 \cdot \psi_{t-1} + \omega_{24} \cdot \psi_{t-24} + \omega_1\omega_{24} \cdot \psi_{t-25}.\end{aligned} \tag{6}$$

模型的创新项 ψ_t 被假设为同方差且服从正态分布，这意味着 $\psi_t \sim N(0, \sigma_\epsilon^2)$。该模型具有 24 小时季节性。由于我们还观察到周季节性，因此我们将输电系统运营商在一周前同一小时的负荷预测$\hat{}$ l_{t-168}作为回归变量纳入模型。

我们基于窗口长度 l_w 对模型进行校准和估计，该窗口包含一年的历史数据。估计后的模型用于递归地（即按小时）预测次日每个小时的数值。由于我们采用自回归时间序列模型，预测小时之前 168 小时至 1 小时的日前负荷预测会作为解释变量纳入两日前预测。然而，这意味着在进行预测时，部分数值尚不可用。这些数值将由基于最新可用观测值递归预测的数值替代。

鉴于提前两天进行预测所伴随的不确定性增加，我们的混合模型纳入了参数密度预测，该预测考虑了负荷水平及其变化的各种情景，并将逐小时提前两天负荷预测作为输入变量。

4.2. 参数密度预测步骤

为考虑提前两天负荷预测值中的不确定性，使用分位数回归平均（QRA）在 5% 和 95% 分位数处计算情景。该方法用于确定预测累积分布函数的分位数，进而可用于构建预测区间。预测区间使用预测累积分布函数的 $(\alpha/2)$阶和 $(1-\alpha/2)$阶分位数 $\alpha \in (0,1)$分别作为区间的下界和上界。QRA 基于分位数回归，旨在对依赖于解释变量 [的实值变量的分位数进行建模（例如参见 Koenker 和 Bassett, 1978]。该方法利用点预测来解释观测值条件分布的 q阶分位数，并设定一个固定的 $q \in (0,1)$。在此，分位数回归使用回归变量向量 $X_t = [1, \hat{l}_t^{2DA}]$，其中包括截距对应的值 1，以及给定时刻 t的负荷提前两日点预测，以计算提前两天负荷预测的 q阶分位数（$Q_{l_t}(q|\cdot)$）（以附加信息为条件）：

$$Q_{l_t}(q|X_t) = X_t\beta_q \tag{7}$$

因此，β_q 是第 q个分位数的参数向量。分位数回归平均通过最小化相应的第 q个分位数的Pinball损失函数来估计 β_q，从而得到第 q个分位数，其定义为

$$\hat{\beta}_q = \underset{\beta}{\operatorname{argmin}}\Big[\sum_{t:l_t \geq X_t\beta}(q - \mathbf{1}_{l_t < X_t\beta})(l_t - X_t\beta)\Big], \tag{8}$$

其中**1** 是指示函数 [参见，例如，Nowotarski 和 Weron, 2015, 2018]。

我们使用提前两日负荷预测$\hat{l}_t^{2DA}$以及一年历史时期内相应的负荷来估计未知参数 β_q。计算出5%和95%分位数后，这些情景覆盖了90%的可能负荷预测值，其中 $\alpha = 0.9$。结合参数密度预测和期望值点预测，该方法为提前两日负荷提供了三种可能的情景：期望场景，以提前两日负荷预测描述；低值场景，以逐小时估计的5%分位数描述；高值场景，以逐小时估计的95%分位数描述。受Grothe [2013] 中的最优集成规则启发（并按照该研究的建议，设定 $\kappa = 2$），我们为期望值赋予权重 2/3，并为每个分位数赋予权重 1/6，以便将这些情景纳入下文所述的*em.power*调度模型。

4.3. 能源系统优化步骤

本节介绍能源系统模型*em.power*调度，该模型在混合模型的能源系统优化步骤中生成批发市场日前价格估计值。模型对欧洲一体化电力部门的主要技术经济方面进行了详细建模，包括市场间的输电限制、热电联产机组的电力生产、储能以及控制功率提供。对于所有纳入考虑的市场区域，我们的模型确定各种发电和储能技术的最优调度决策、跨境输电容量的最有效利用方式，以及短期系统边际成本[5]，；这些成本决定了日前市场在小时分辨率下的价格估计值。

由于我们的研究聚焦于日前价格预测，因此能源系统模型旨在反映交付前一天市场参与者可获得的信息水平和质量。该模型被表述为一个以系统总成本最小化为目标的线性优化问题。为确保高度复杂的系统采用线性形式，我们构建容量集群，并按照第3节所述对其进行参数化。在每个容量集群中，机组可以启动，电力也可以按边际增量生产 [参见Muesgens，2006b]。这种方法有两个主要优势。首先，它降低了计算需求。其次，该问题在每一点上都是可微的，需求约束的对偶变量可以解释为批发市场价格估计值。此外，对于我们的研究目的而言，对大型能源系统的建模仍具有相当高的准确性 [Muesgens 和 Neuhoff， 2006]。

我们通过模型的两个特征来实现由此产生的不完美预测。首先，我们采用滚动窗口方法，反复求解三天（$d \in D = \{d, d+1, d+2\}$），如图3所示。在这种设定下，目标日的24小时对应于预测期限的第二天（$d+1$）。这遵循 EPEX现货市场的组织方式：交付前一天中午12点确定24个逐小时的日前价格（d）。除目标日 d+1外，我们还纳入前一天（d）和后一天（$d+2$）。通过在滚动窗口中考虑三天，我们减少了起始值和结束值带来的问题，尤其是由发电厂启动和抽水蓄能电站引起的问题。其次，我们考虑关键参数两日前估计值的不确定性增加。第 d 天的参数完全已知，而第 $d+1$ 天的预测由欧洲电力输电系统运营商网络（Entso-e）提供；关键输入参数的实际取值在第 $d+2$天具有更高的不确定性。因此，我们仅在第 $d+2$ 天采用概率预测区间，如图3所示。所有其他日期均通过一个情景表示。随后，使用市场参与者做出决策时可获得的数据，反复求解由此得到的随机滚动窗口。在每次模型运行中，我们提取“目标日”——即日前——24小时的信息。

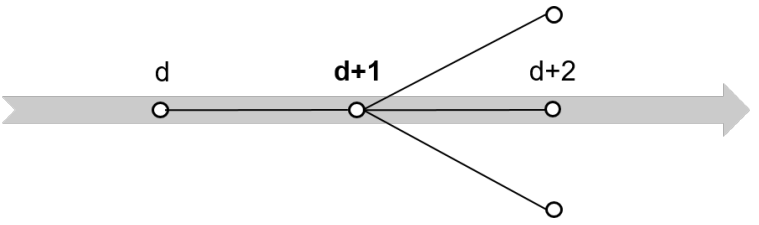


图 3： 滚动窗口示意图。

优化问题每天都会重复求解。为方便读者，我们在附录中提供了符号表。请注意，所有内生变量均以大写字母表示，所有外生参数均以小写字母表示。

公式（9）中的目标函数旨在最小化系统总成本（TC），该成本由滚动窗口 d*覆盖的三天内所有运营成本的期望值（OC）构成。我们的实证研究考虑了三种情景 s，并假设它们出现的可能性相同。

$$\min TC = \mathbb{E}_s[OC(s)] \tag{9}$$

[5]严格来说，需求约束的对偶变量即为短期系统边际成本，也称为“影子价格”。

公式（10）中的运营成本包含发电机组承担的所有短期成本。我们纳入满负荷运行成本（$vc^{FL}_{i\,n\,d}$、、、h）、部分负荷运行机组的额外成本（$vc^{ML}_{i\,n\,d}$、、、h − $vc^{FL}_{i\,n\,d}$、、、h）以及启动成本（sc_{i}、n、d、h）。请注意，我们采用线性机组组合模型，且所有机组的发电量必须至少达到一定的最低出力水平（见公式（12）–（16））。此外，我们还计入切负荷成本（$voll$）以及因削减可再生能源而产生的罚款（$curtc$），如第 3节所述。

$$
\begin{aligned}
OC_s = & \sum_{i,n,d,h} G_{i,n,d,h,s} \cdot vc^{FL}_{i,n,d,h} + \sum_{i,n,d,h} SU_{i,n,d,h,s} \cdot sc_{i,n,d,h} \\
& + \sum_{i,n,d,h} (P^{on}_{i,n,d,h,s} - G_{i,n,d,h,s}) \cdot (vc^{ML}_{i,n,d,h} - vc^{FL}_{i,n,d,h}) \cdot g^{min}_i / (1 - g^{min}_i) \\
& + \sum_{stl,n,d,h} G_{stl,n,d,h,s} \cdot wv_{stl,n,d,h} - \sum_{stl,n,d,h} CL_{stl,n,d,h,s} \cdot wv_{stl,n,d,h} \\
& + \sum_{hr,n,d,h} G_{hr,n,d,h,s} \cdot wv_{hr,n,d,h} + \sum_{n,d,h} SHED_{n,d,h,s} \cdot voll \\
& + \sum_{res,n,d,h} CURT_{res,n,d,h,s} \cdot curtc
\end{aligned}
\tag{10}
$$

由于我们采用滚动窗口应用模型，每次模型运行都会考虑三天 d，每天24小时，即每次日模型运行涵盖72小时。对于按日循环运行的储能单元而言，在目标日前后各额外建模一天似乎是合理的。然而，储能单元（包括抽水蓄能电站和不带泵的季节性储能机组）的储能循环超过三天。因此，抽水蓄能电站被划分为低容量能量比储能和高容量能量比储能：前者的储能循环超过三天，后者则在三天预测期限内运行一个或多个储能循环。后者的调度由模型内生决定。低容量能量比抽水蓄能电站被赋予水价值（$wv_{stl,\ n,\ d,\ h}$），该水价值作为发电（$G_{stl,\ n,\ d,\ h,\ s}$）和用电量（$CL_{stl,\ n,\ d,\ h,\ s}$）的可变成本系数予以实施。

除抽水蓄能电站外，水库具有自然水流入，且未安装水泵。然而，用于发电的水量预算受到季节性流入水量的限制。我们还采用水价值（$wv_{hr,\ n,\ d,\ h}$），以计入利用水库发电的机会成本（$G_{hr,\ n,\ d,\ h,\ s}$）。

市场出清由公式（11）保证。在给定滚动窗口的全部 72 小时内，需求（$l_{n,d,h,s}$）必须等于发电量（$G_{i,n,d,h,s}$）、切负荷（$SHED_{n,d,h,s}$）和电力进口（$FLOW_{nn,n,d,h,s}$）之和，减去中期储能（$CM_{stm,n,d,h,s}$）和长期储能（$CL_{stl,n,d,h,s}$）的用电量，再加上电力出口（$FLOW_{n,nn,d,h,s}$）。需求约束公式（11）的对偶变量用作小时级日前批发电价估算值。

$$
\begin{aligned}
l_{n,d,h,s} = & \sum_{i} G_{i,n,d,h,s} - \sum_{stm \subset I} CM_{stm,n,d,h,s} - \sum_{stl \subset I} CL_{stl,n,d,h,s} \\
& + \sum_{nn} (FLOW_{nn,n,d,h,s} - FLOW_{n,nn,d,h,s}) + SHED_{n,d,h,s} \\
& \forall n, nn \in N, d \in D, h \in H, s \in S
\end{aligned}
\tag{11}
$$

请注意，我们将第4.1节中为德国在 d 和 d+1提出的改进点预测应用于德国，因为我们关注的是德国市场的电价预测。对于所有其他市场，我们采用 Entso-e 的原始预测。对于德国的 d+2，我们采用第4.2节中介绍的概率负荷预测。对于其余市场，我们使用前一周的实际实现值作为一种较为朴素的估算值。

容量集群的发电量受上下界限制。上界在式（12）中进行了形式化。该上界确保集群中的发电量不超过运行容量（$P^{on}_{i\,n}$、、d、h、s）。运行容量的发电潜力还受到正向调节功率供给备用容量的进一步限制（PCR_{i}、n、bp、s和 $SCR^{pos}_{i\,n}$、、bs、s）。下界在式（13）中给出。该式表明，运行容量必须以最小功率水平运行，其中包括为负向调节功率供给预留的容量（PCR_{i}、n、bp 以及

$SCR^{neg}_{i\,n}$ 、、$_{bs、\,s}$）。需要注意的是，德国的一次控制功率提供（$PCR_{i、\,n、\,bp、\,s}$）是对称的（即一个机组必须同时提供正向和负向一次控制功率）。对于二次控制电力，引入了不同的正向和负向控制电力产品。出于两个原因，我们没有在模型中纳入分钟备用要求。首先，快速响应机组（例如水电、开式循环燃气轮机）可以在不参与调度的情况下启动，以提供正向分钟备用。其次，除模型中包含的发电厂之外，多个市场参与者也能够提供正向和负向备用容量（例如需求灵活性、P2X机组、应急发电机）。属于竞价块的小时通过 bp 映射至一次控制功率，通过 bs映射至二次控制电力。

$$\begin{aligned} &G_{i,n,d,h,s} \le P^{on}_{i,n,d,h,s} - PCR_{i,n,bp|h\in bp,s} - SCR^{pos}_{i,n,bs|h\in bs,s} \\ &\forall bp \in BP, bs \in BS, i \in I, n \in N, d \in D, h \in H, s \in S \end{aligned} \tag{12}$$

$$\begin{aligned} &P^{on}_{i,n,d,h,s} \cdot g^{min}_{i} + PCR_{i,n,bp|h\in bp,s} + SCR^{neg}_{i,n,bs|t\in bs,s} \le G_{i,n,d,h,s} \\ &\quad \forall bp \in BP, bs \in BS, i \in I, n \in N, d \in D, h \in H, s \in S \end{aligned} \tag{13}$$

装机容量限制了电力系统的运行容量（$cap_{i,\ n,\ d,\ h}$），这与可用率因子（$af_{i,\ n,\ d,\ h}$）或发电厂停运（$out_{i,\ n,\ d,\ h}$）有关，如方程（14）所示。对于火力发电容量，我们采用发电厂逐小时停运数据。可再生能源的可用率因子按小时提供，而水电机组的月度可用率因子按月提供。

$$P^{on}_{i,n,d,h,s} \le cap_{i,n,d,h} \cdot af_{i,n} - out_{i,n,d,h} \qquad \forall i \in I, n \in N, h \in H, d \in D, s \in S \tag{14}$$

方程（15）跟踪使运行容量从一小时增加到下一小时的启动活动（$SU_{i,\ n,\ d,\ h,\ s}$）。由于非负条件，启动量只能为正或为零。方程（16）跟踪从一天最后一小时到次日第一小时的启动活动。

$$P^{on}_{i,n,d,h,s} - P^{on}_{i,n,d,h-1,s} \le SU_{i,n,d,h,s} \qquad \forall i \in I, n \in N, h \in H, d \in D, s \in S \tag{15}$$

$$P^{on}_{i,n,d,hfirst,s} - P^{on}_{i,n,d-1,hlast,s} \le SU_{i,n,d,hfirst,s} \qquad \forall i \in I, n \in N, d \in D, s \in S \tag{16}$$

间歇性可再生能源的可用上网电量与其实际发电量之差定义为可再生能源弃电，如方程（17）所示:

$$\begin{aligned} &cap_{res,n,d,h} \cdot pf_{res,n,d,h} = G_{res,n,d,h,s} + CURT_{res,n,d,h,s} \\ &\forall n \in N, res \in I, d \in D, h \in H, s \in S. \end{aligned} \tag{17}$$

一些发电厂同时活跃于热力市场和电力市场。因此，该模型针对这些机组在电力市场中实施了一个随时间变化的必运行条件（例如，冬季因空间供暖需求而提高）。根据小时热需求，式（18）指出，热电联产机组的输出至少等于与热量生产相关的发电量（$chp_{i,n,d,h}$）：

$$chp_{i,n,d,h} \le G_{i,n,d,h,s} \qquad \forall i \in I, n \in N, d \in D, h \in H, s \in S. \tag{18}$$

公式 19 通过净传输容量约束跨境电力传输（$FLOW_{n,\ nn,\ d,\ h,\ s}$）（$ntc_{n,\ nn,\ d,\ h}$）：

$$FLOW_{n,nn,d,h,s} \leq ntc_{n,nn,d,h} \qquad \forall n, nn \in N, d \in D, h \in H, s \in S. \tag{19}$$

式（20）描述了中期储能单元的储能水平状态。发电会使储能水平下降（$G_{stm,\ n,\ d,\ h}$），充电则会使其上升（$ST^{in}_{stm,\ n,\ d,\ h,\ s}$）。整个储能循环效率（$\eta_{stm}$）计入充电过程。式（21）确保储能机制在两天之间正常运行：

$$\begin{aligned} &SL_{stm,n,d,h,s} = SL_{stm,n,d,h-1,s} - G_{stm,n,d,h,s} + CM_{stm,n,d,h,s} \cdot \eta_{stm} \\ &\forall stm \in I, n \in N, d \in D, h \in H, s \in S, \end{aligned} \tag{20}$$

$$\begin{aligned} SL_{stm,n,d,hfirst,s} =& SL_{stm,n,d-1,hlast,s} - G_{stm,n,d,hfirst,s} \\ &+ CM_{stm,n,d,hfirst,s} \cdot \eta_{stm} \\ \forall stm \in I, n \in N, d \in D, s \in S. \end{aligned} \tag{21}$$

高容量能量储存能力（$SL_{stm,\ n,\ d,\ h,\ s}$）对于容量能量比较高的储能单元，其最大值定义为最大已安装涡轮机容量除以容量能量比（cer），如式（22）所示：

$$SL_{stm,n,d,h,s} \leq cap_{stm,n,d,h} \cdot cer \qquad \forall stm \in I, n \in N, d \in D, h \in H, s \in S. \tag{22}$$

式（23）同时限制涡轮机容量和泵送容量，并假设泵送容量比涡轮机容量低 10%：

$$\begin{aligned} &G_{stm,n,d,h,s} + 1.1 \cdot CM_{stm,n,d,h,s} \leq cap_{stm,n,d,h} \\ &\forall stm \in I, n \in N, d \in D, h \in H, s \in S. \end{aligned} \tag{23}$$

在每次模型运行开始和结束时，所有容量能量比较高的储能单元，其储能水平必须至少达到 30%：

$$SL_{stm,n,dfirst,hfirst,s} = 0.3 \cdot cap_{stm,n,d,h} \qquad \forall stm \in I, n \in N, s \in S, \tag{24}$$

$$SL_{stm,n,dlast,hlast,s} = 0.3 \cdot cap_{stm,n,dlast,hlast} \qquad \forall stm \in I, n \in N, s \in S. \tag{25}$$

容量能量比较低的储能电站不受储能机制约束。然而，这些单元的发电量和用电量受其装机容量限制，具体如下：

$$G_{stl,n,d,h,s} + CL_{stl,n,d,h,s} \leq cap_{stl,n,d,h} \qquad \forall n \in N, stl \in I, d \in D, h \in H, s \in S. \tag{26}$$

公式（27）、（28）和（29）确保提供一次控制电力、正二次控制电力和负二次控制功率。

$$\sum_i PCR_{i,n,bp,s} = pr_n \qquad \forall bp \in BP, n \in N, s \in S \tag{27}$$

$$\sum_i SCR^{pos}_{i,n,bs,s} = sr^{pos}_n \qquad \forall bs \in BS, n \in N, s \in S \tag{28}$$

$$\sum_i SCR^{neg}_{i,n,bs,s} = sr^{neg}_n \qquad \forall bs \in BS, n \in N, s \in S \tag{29}$$

非负约束确保各个变量不呈现负值，其形式如下：

$$\begin{aligned} 0 \leq & CL_{stl,n,d,h,s}, CM_{stm,n,d,h,s}, CURT_{res,n,d,h,s}, \\ & G_{i,n,d,h,s}, FLOW_{n,nn,d,h,s}, P_{i,n,d,h,s}, PCR_{i,n,bp,s}, SCR^{neg}_{i,n,bs,s}, SCR^{pos}_{i,n,bs,s}, \\ & SHED_{n,d,h,s}, SL_{stm,n,d,h,s}, SU_{i,n,d,h,s} \\ & \forall n, nn \in N, bp \in BS, bs \in BS, i \in I, n \in N, h \in H, d \in D, s \in S. \end{aligned} \tag{30}$$

4.4. 随机数据后处理步骤

在此步骤中，我们使用随机后处理技术来细化技术经济能源系统模型生成的价格估计器。具体而言，我们对能源系统优化步骤获得的日前价格估计值的误差 ε_t 进行预测，方法可以是基于时间序列的点预测 $\hat{\varepsilon}_t$，也可以通过推断预测误差的分布函数来生成日前价格概率预测。我们将可再生能源上网电量和天气等外生变量，以及预测误差本身的滞后项纳入对 ε_t的时间序列预测中。

我们从 时刻 t的改进点预测 $\hat{P}_t^*$ 开始，该预测由以下方程给出 ：

$$\hat{P}_t^* = \hat{P}_t + \hat{\varepsilon}_t, \tag{31}$$

其中， $\hat{P}_t$ 是上一步的电价预测， $\hat{\varepsilon}_t$ 是我们模型预测的电价预测误差。因此， $\hat{P}^*$构成了改进后的价格预测，其中我们针对误差中随机但可预测的结构，对上一步的预测进行了调整。

我们采用单变量和多变量两种模型框架，因为以往研究已证明这种方法很有用Ziel and Weron [2018]。在单变量框架中，我们将预测误差时间序列视为一个小时分辨率的高频时间序列。在多变量框架中，我们将时间序列拆分为 24 个独立的时间序列，每个小时对应一个，使其具有日分辨率。

对于后处理设置，下标 h, d 表示第 d 天的第 1 至第 24 小时，其中 d 表示连续的天数。因此，例如，$P_{1,1}$ 表示所研究时段第一天第一小时的实际日前电价，而$\varepsilon_{5,432}$ 表示第 432 天第五小时的电价估计器的误差。这样的设置最为合适，因为它能让我们同时观察次日 24 个小时的电价实现值。请注意，如果 $h-1$小于或等于零，我们就需要向前回溯一天。同样，如果 $h+1$ 大于 24，我们就需要向前移一天。

我们为单变量框架选择了标准计量经济学时间序列模型。该模型由内生变量（即具有移动平均结构的自回归变量）和外生变量组成，所有这些变量均整合到式（32）给出的回归模型中。为处理价格预测误差时间序列中包含的若干季节性结构 $\varepsilon_{h,\ d}$，我们使用向前回溯的第一和第二个观测值 $\varepsilon_{h-1,\ d}$，$\varepsilon_{h-2,\ d}$，以及估计模型的第一个滞后误差 $\psi_{h-1,\ d}$。此外，我们还使用前一天（日结构）$\varepsilon_{h,\ d-1}$ 和前一周（周结构）$\varepsilon_{h,\ d-7}$ 的观测值作为内生解释变量。考虑到日效应，我们纳入前一天的最小和最大预测误差 $\varepsilon_{min,\ d-1}$，$\varepsilon_{max,\ d-1}$。为考虑预测误差的强烈影响

在公共假日期间，我们使用公共假日虚拟变量 $hol_{h,d}$ 作为另一个因素。此外，我们纳入逐小时风电预测 $X_{h,d}$。

$$\begin{aligned}\varepsilon_{h,d} =& \phi_0 + \phi_1 \cdot \varepsilon_{h-1,d} + \phi_2 \cdot \varepsilon_{h-2,d} + \phi_3 \cdot \varepsilon_{h,d-1} + \phi_4 \cdot \varepsilon_{h,d-7} + \phi_5 \cdot \psi_{h-1,d} \\ &+ \omega_1 \cdot \varepsilon_{min,d-1} + \omega_2 \cdot \varepsilon_{max,d-1} + \omega_3 \cdot hol_{h,d} + \omega_4 \cdot X_{h,d} \\ &+ \psi_{h,d}\end{aligned} \tag{32}$$

与单变量框架一样，我们在多变量框架中使用广为人知的 ARX模型时间序列模型。不过，自回归部分指的是前几天同一时刻的数值。内生变量 $\varepsilon_{h,\ d-1}$ 和 $\varepsilon_{h,\ d-7}$分别是前一天和前七天同一时刻的预测误差。外生变量与单变量框架中的相同：前一天预测误差的最小值和最大值、一个表示公共假日的虚拟变量，以及逐小时风电预测。因此，多变量框架的模型由下式给出：

$$\begin{aligned}\varepsilon_{h,d} =& \phi_0 + \phi_1 \cdot \varepsilon_{h,d-1} + \phi_2 \cdot \varepsilon_{h,d-7} \\ &+ \omega_1 \cdot \varepsilon_{min,d-1} + \omega_2 \cdot \varepsilon_{max,d-1} + \omega_3 \cdot D_{h,d} + \omega_4 \cdot X_{h,d} \\ &+ \psi_{h,d},\end{aligned} \tag{33}$$

其中，ϕ_i，ω_i 表示需要估计的系数。在两种框架中，创新项 ψ 均假设为同方差且服从正态分布，这意味着 $\psi_{h,\ d} \sim N(0, \sigma_\epsilon^2)$。

由于我们采用自回归时间序列模型，因此需要使用预测时点前最后几个小时的日前现货价格作为解释变量。在多变量框架中，由于每个小时分别建模，因此每次只预测未来一个时点。在单变量框架中，则预测未来的 24 个值，并递归预测次日各小时的数值（即逐小时进行预测）。对于不可用的变量，则使用基于最近可用观测值递归预测得到的变量替代。

与文献中先前证明有效的方法一致 [例如，参见Marcjasz 等人， 2018], 我们通过使用不同的窗口长度来估计模型设定并避免随机选择，从而改变所提出的后处理模型。我们将校准窗口设为 44、48 和 52 周。使用三种窗口长度和两种模型框架后，我们得到六个单个子模型。这些子模型用于预测h次日 d各小时的数值，并将它们记作$\hat{P}_{h\,d}^{uv44}$,, $\hat{P}_{h\,d}^{uv48}$,, $\hat{P}_{h\,d}^{uv52}$,, $\hat{P}_{h\,d}^{mv44}$,, $\hat{P}_{h\,d}^{mv48}$,, $\hat{P}_{h\,d}^{mv52}$;

我们通过对六个价格取算术平均值，得到最终改进的点预测 $\hat{P}_t^*$。尽管使用看似更复杂的方法可能颇具吸引力，例如根据历史数据通过线性回归计算最优权重，但这些方法得到的预测器均方根误差（RMSE）和平均绝对误差（MAE）更高，即使我们使用了向前看的滚动窗口也是如此。这主要是因为使用此类方法会引入额外的估计噪声，并可能导致效率低下；例如，金融文献中的DeMiguel 等人 [2009]在相关讨论中也指出了这一点。因此，我们仍采用较为简单但更稳健的方法，对六个单独的预测取平均值。

现在，我们继续生成日前价格概率预测。为此，我们使用单个子模型生成的六个预测来估计日前价格的累积分布函数 $F_{P_{h,\ d}}$^估计得到的函数 $F_{P_{h,\ d}}$ 作为我们对次日价格的概率预测。我们用其分位数来表示该^分布 $F_{P_{h,\ d}}$。具体而言，我们采用分位数回归来对日前价格累积分布函数的条件 q 阶分位数进行建模，其中 q 是介于 0 和 1 之间的值。该建模过程利用六个单独的点预测和以下方程完成：

$$Q_{P_{h,d}}(q|X_{h,d}) = X_{h,d}\beta_q, \tag{34}$$

其中 $X_{h,\ d} = [1, \hat{P}^{uv44}_{h,d}, \hat{P}^{uv48}_{h,d}, \hat{P}^{uv52}_{h,d}, \hat{P}^{mv44}_{h,d}, \hat{P}^{mv48}_{h,d}, \hat{P}^{mv52}_{h,d}]$ 是回归变量向量，其中包含截距对应的值 1，以及时间 h，d日前电价的六个单独点预测。 β_q同样通过最小化 Pinball损失函数进行估计。为了确定预测分布，我们对该分布的多个分位数进行预测。

回归变量的系数通过为期一年的校准窗口进行估计，并区分高峰时段和非高峰时段。高峰时段定义为周一至周五上午 8 点至晚上 8 点之间的时段，而非高峰时段则包括其余所有时段。之所以作此区分，是因为高峰时段具有较高的电力需求，因此日前价格通常也较高。对于每个分位数 q，我们估计两个参数向量 β_q（见公式（8））。事实证明，为所有日内时段估计一个参数向量 β_q 的准确性较低。如有需要，可提供有关此事项的更多信息。

总之，我们的混合模型包括以下步骤。为了预测次日情况，我们首先在随机数据预处理步骤中改进输电系统运营商的日前负荷预测，并预测提前两日负荷预测。随后，我们在参数密度预测步骤中计算提前两日负荷情景，并将其纳入能源系统优化步骤中的 *em.power*调度，以生成日前现货市场的第一批价格估计器。最后，在最终的随机后处理步骤中，我们使用随机方法改进这些价格估计器，并进行概率电价预测。对于观测期内的所有时间点，该流程按日持续重复进行。混合模型采用的滚动窗口方法意味着我们始终使用最新可用数据。

5. 混合模型结果

本节介绍我们的混合模型对德国2016年1月至2020年12月的电价预测。由于日前市场以小时分辨率组织交易，我们的混合模型会针对次日的每个小时计算点预测和概率预测。作为本文的核心内容，我们展示模型的总体结果（即点价格预测和概率价格预测），并定性地详细说明其在相关文献中的定位。

我们首先将混合模型的点预测与文献中的结果进行比较。在2016年至2020年的五年间，我们模型的年均均方根误差为7。38 =欧分/兆瓦时，平均绝对误差为4。60 =欧分/兆瓦时，与以往研究的结果相当。由Ziel and Weron [2018] 开发的专家模型预测了2012年至2016年的电价，整体平均绝对误差为5。01 =欧分/兆瓦时。Maciejowska等人采用带外生变量的自回归模型，Maciejowska 等 [2021] 在2016年至2019年间取得了8。43 =欧分/兆瓦时的均方根误差和5。92 =欧分/兆瓦时的平均绝对误差。在同一时期，Qussous等人 [2022] 取得了11。21 =欧分/兆瓦时的均方根误差和7。89 =欧分/兆瓦时。他们提出了一种基于代理的电力市场仿真模型，该模型采用基于规则的竞价策略，因此属于非均衡导向的技术经济市场模型，旨在重现日前电价。由于他们提供了详尽的误差度量信息，我们可以利用这些模型进行详细比较。不过，由于这些研究的时间段与我们的观测期部分重叠但并不完全一致，我们还将其与Lago等人 [2021]开发的LEAR模型进行详细比较。LEAR模型的代码可免费获取，因此我们可以将其扩展至纳入截至2020年的数据。此外，该模型的日前电价预测在文献中准确度名列前茅，其作者也是价格预测领域的领军学者。因此，我们可以逐年将其结果与我们的结果进行比较。表2 展示了本文提出的混合模型、基于代理的市场仿真模型和LEAR模型的预测准确度，并列出了年度均方根误差和平均绝对误差。可以看出，基于代理的模型误差最高，这可归因于技术经济模型在短期预测方面普遍面临的困难（em.power调度模型未经进一步后处理的结果汇总于第*em.power*调度节，并指向相同结论）。相比之下，LEAR模型与本文提出的混合模型误差度量相近，差距并不大。尽管LEAR模型更常占优，但其优势有限，因为它是一个主要为生成价格预测而设计的统计模型。相比之下，混合模型的预测涵盖了能源系统模型所表征的整个市场状态，其中包括

[6]更确切地说，价格数据涵盖2016年1月至2018年9月的德国-奥地利-卢森堡竞价区，以及2018年10月至2020年12月的德国-卢森堡竞价区。

表 2通过所提出的混合模型和 LEAR 模型进行的日前电价预测的均方根误差和平均绝对误差，单位为[欧分/兆瓦时]

=

	RMSE			MAE		
	混合式	Agent	LEAR	混合	Agent	LEAR
所有年份	7.38	11.21	7.24	4.60	7.89	4.38
2016	5.82	8.83	5.55	3.48	6.54	3.30
2017	8.79	13.01	7.91	5.25	9.44	4.56
2018	7.28	11.69	6.96	5.07	8.88	4.84
2019	7.05	10.91	7.95	4.43	6.69	4.53
2020	7.63		7.54	4.77		4.65

市场参与者感兴趣的其他参数信息（例如 CO_2 排放、国际电力交换和电厂利用率）。[7]

对混合模型预测误差的更详细评估见图表4 和5，分别展示了基荷、高峰和非高峰时段等不同标准下的均方根误差，以及实际日前价格分位数对应时段的均方根误差。同样，高峰时段指周一至周五上午8点至晚上8点之间的时段；其余时段均为非高峰时段。基荷时段指一天中的所有时段，不论星期几或一天中的哪个时段。

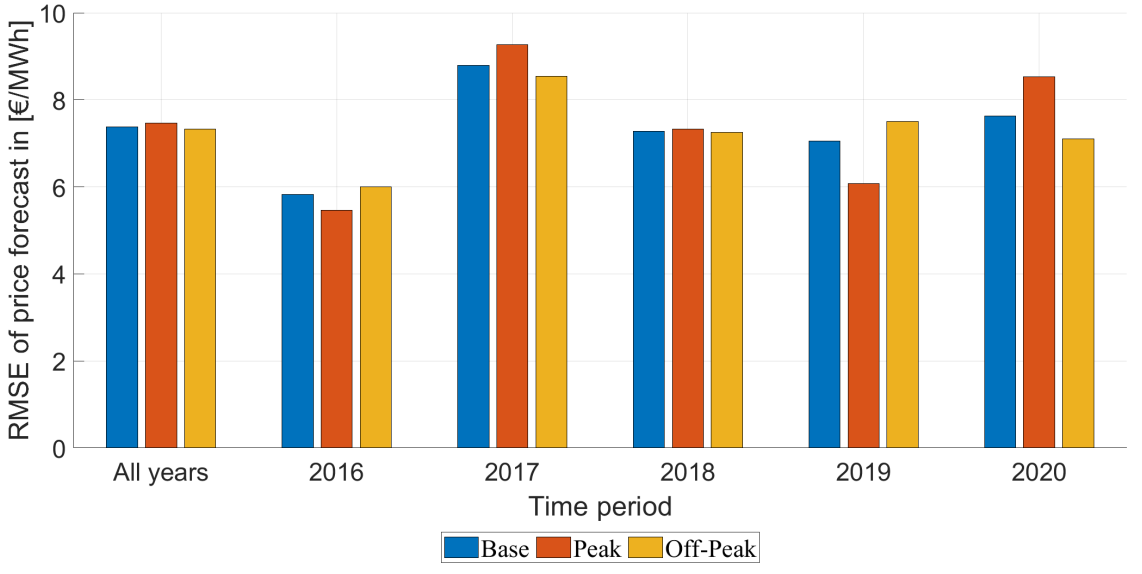


图4：基荷、高峰和非高峰时段的均方根误差。

尽管7。47 =欧分/兆瓦时的高峰时段误差与7。33 =欧分/兆瓦时的非高峰时段误差在整个期间内看起来相当相似，但图5 提供了更多信息。我们将预测时段分为五组，按照实际日前价格的每个置信区间以20%的步长呈现各组时段。因此，我们计算并评估每组的均方根误差。图表显示，价格预测误差与价格水平之间存在一定关系。多年来，价格最低20%和最高20%的时段始终被确定为均方根误差最高的时段。根据经济理论，这一发现可以通过启动成本及其对小时价格的影响来解释。假设完全预见，Kuntz 和 Muesgens [2007] 表明，在一个负荷周期中，启动成本仅会在需求最高的时段加到燃料成本中，因为该时段必须启动额外产能，而其他时段则不需要。在需求最低的时段，启动成本会从可变生产成本中扣除，因为

[7]请注意，尽管本文仅关注价格，但正如文献综述中所讨论的，该模型也能为其他因素提供参考。

发电厂如果获准在整个时段持续运行，就能节省重新启动成本。相比之下，启动成本不会影响负荷周期中的其他任何时段的价格。*em.power*调度 模型在根据需求约束的影子价格确定批发电价时遵循这一经济理论。然而，在现实中，日前市场的竞价者面临着诸多不确定性，包括哪个时段的剩余需求最高或最低，以及当天的启动成本规模。尽管不确定性始终存在，但当除了燃料成本之外还需要考虑启动成本时，其影响可能会更大，从而增加最高和最低价格时段附近的价格波动性。因此，负价格峰值和正价格峰值比中间价格水平更难捕捉和预测。需要注意的是，这些市场条件下不确定性的增加会影响所有点预测模型。LEAR模型的相应图表呈现出非常相似的模式，可参见附录中的图表A.1 和A.2。

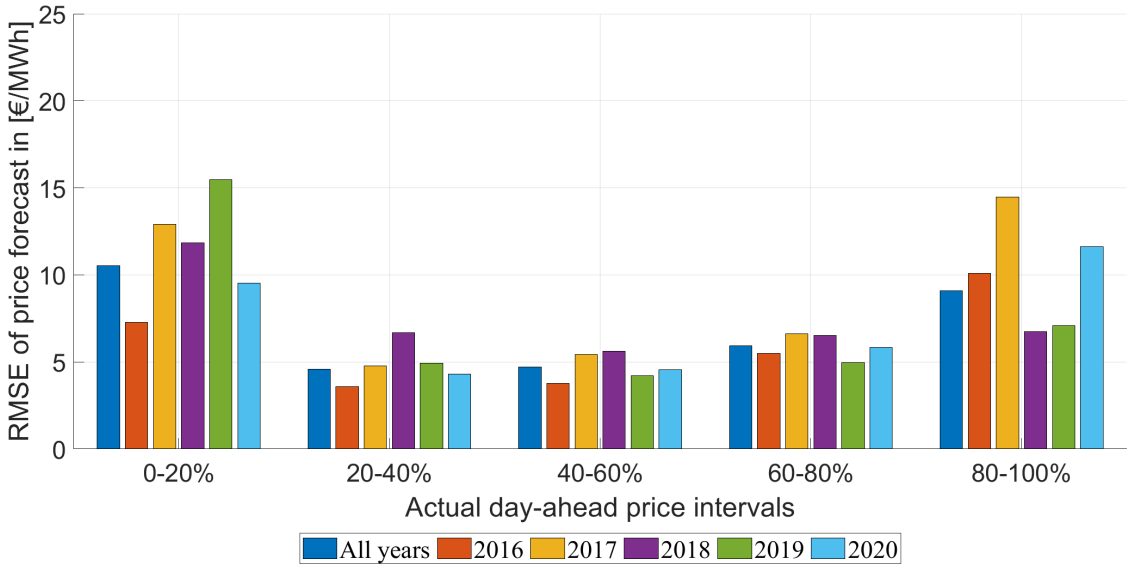


图5：不同日前价格分位数下各小时的均方根误差。

现在我们转向对概率预测的分析。我们使用以 5% 为步长的分位数估计，表示每个预测小时的日前价格预测分布。随后，对不同分位数水平进行分位数估计，即可计算预测区间，从而确定日前电价处于某一特定范围内的概率。

首先需要检查概率预测是否经过良好校准 [例如，参见Gneiting 等人，2007], 也就是说，检查预测概率与观测频率是否一致。因此，图6 展示了日前价格高于或低于预测分位数界限的所有小时的实际频率。如果预测概率与目标变量随时间变化的观测频率相匹配，则认为预测经过了校准。因此，在实验室环境中，完全校准的预测应对应于均匀分布。下方的直方图显示，向两端的频率略有增加，但总体上很好地近似于由随机抽取的数字生成的均匀分布。这为我们的混合模型提供了质量检查，并确保预测概率反映日前价格的真实概率。

由于高峰时段的日前价格通常高于非高峰时段，我们的混合模型针对每个分位数分别使用两组 QRA 估计来预测分位数，一组对应高峰时段，另一组对应非高峰时段。因此，应分别在两个不相交子集中评估概率预测的校准情况。图7 左侧展示了高峰时段分位数的校准情况，右侧展示了非高峰时段的校准情况。在非高峰时段，两侧最外侧的 5% 分位数（即价格水平较高和较低时）的出现频率更高。在高峰时段，超过分位数值的小时数更多，尤其是在中位数右侧。例如，有 6.3% 的小时数超过了 95%分位数值。较低的 5%分位数出现频率更高，最低价格可归因于可再生能源发电量较高所产生的降价影响。这与高峰时段所占比例相符

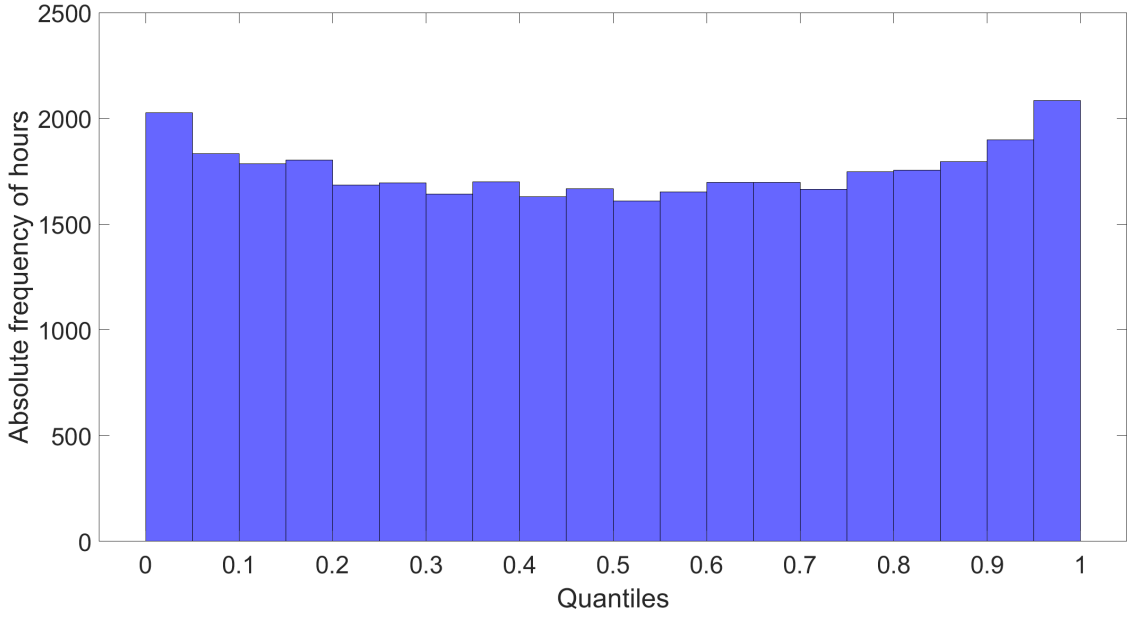


图6： 预测期（2017年至2020年）估计的分位数中包含的实际价格数量。

显示出对电力的高需求，同时由于较高的太阳辐射或风速，可再生能源的馈入量也很高，因此供应量较大。

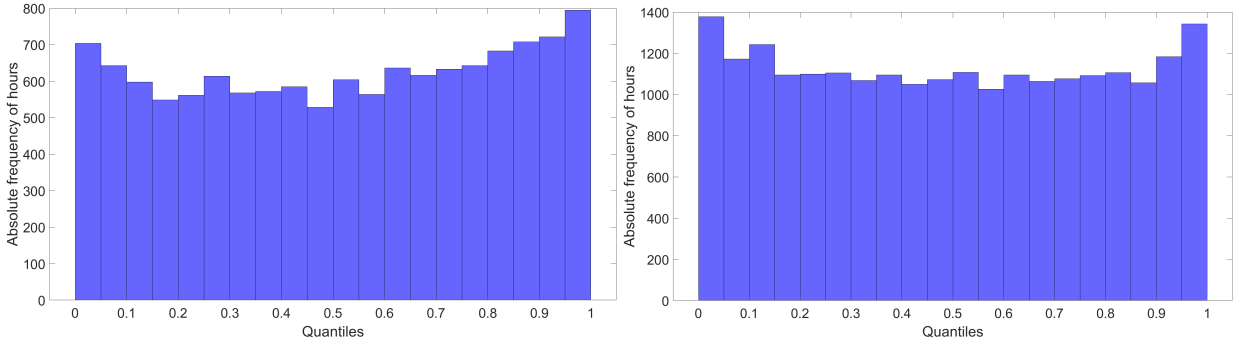


图7： 2017年至2020年预测期内估计的分位数中包含的实际价格数量，分别按高峰时段（左）和非高峰时段（右）列示。

对预测分位数进行校准分析，验证了计算所得概率预测的有效性。接下来，我们评估预测区间宽度的逐小时分布，以更精确、量化地了解市场不确定性及其对预测准确度的影响。此外，我们利用概率预测计算负价格出现的概率。因此，我们探讨了在交易策略和电厂部署规划中纳入概率预测的经济意义。

在高峰时段，日前价格的不确定性和波动性高于非高峰时段。这可以由高负荷水平下更陡峭的merit order、高峰时段启动的发生，以及可再生能源来源发电量的不确定性来解释（尤其是光伏电站，它们在高峰时段的发电量更高）。我们的混合模型通过概率预测捕捉到了这种不确定性。图 8展示了每日各小时估计的95%分位数与估计的5%分位数之间价差分布的箱线图（即90%预测区间的宽度）。从上午8点到晚上8点（白天），中位数价差更宽，离群值也高于夜间，反映出更高的不确定性和更广泛的价格范围。尽管更宽的价差

涵盖90%潜在价格的预测区间表明，这些时段比夜间时段更难预测，但混合模型对这些时段的预测准确度与夜间时段一样高。图 4 显示，在整个期间内，高峰时段和非高峰时段具有相当的点预测准确性。此外，概率预测有效地反映了市场的不确定性。

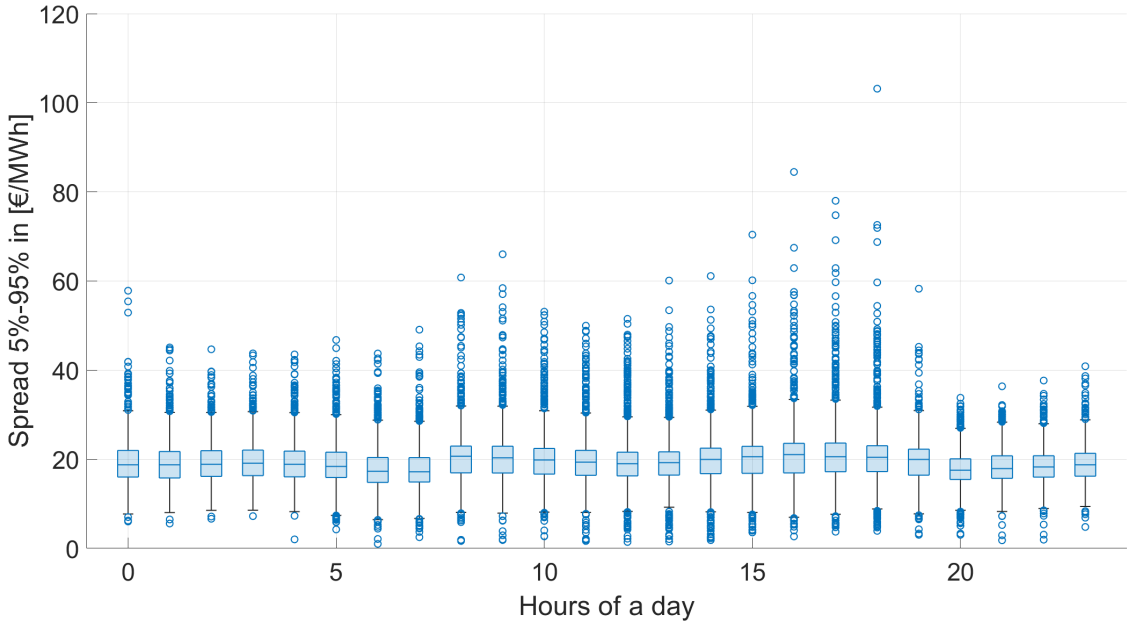


图8：一天中每个小时的90%预测区间箱线图。

图9展示了一周中每个小时的90%预测区间平均宽度，说明一周中各个小时的日前电价预测难易程度。价差越大，表明不确定性越高、可预测性越低，因为价格可能取值的范围更广。如前所述，白天时段的区间通常比夜间更宽。不过，还可以观察到更细致的规律。

在工作日（周一至周五），预测区间的价差呈现出明显的日变化规律，形成三条凹形曲线，局部最大值分别出现在凌晨3点、上午8点和下午4点。午夜至凌晨7点，区间相对较窄。上午7点至8点，区间骤然扩大，给电价预测带来重大挑战，尤其是在下午4点至6点之间。在周末（周六和周日），区间宽度在各个时段的分布更加均匀，但总体仍然较高，并且在晚上7点至9点之间仍有明显下降。因此，预测日前电价时，除了所有其他可能影响市场动态的因素外，还需仔细考虑一天中的时段和一周中的日期。

图 10 展示了概率预测如何体现不确定性。图中显示了下限和上限的变化，以及预测的期望价格（价格的点预测）。所示上下限分别对应估计的 5%分位数和 95%分位数，表示可能结果的范围。更准确地说，实际日前电价有 90% 的概率落在该范围内。借助密度预测和预测边界，我们可以做出有依据的风险估计并提出概率陈述，这有助于在不确定性下进行决策。

图 11展示了根据密度预测估计的每周各时段负价格发生概率。值得注意的是，在某些情况下可能出现负价格，例如，可再生能源发电量增加或必须运行容量占比较大（如热电联产）时，可能迫使常规发电机组停机，而当再次需要这些容量时便会产生启动成本。可再生能源的优先次序效应以及可再生能源监管，尤其是电力批发市场之外的发电补贴，可能会加剧这一情况。

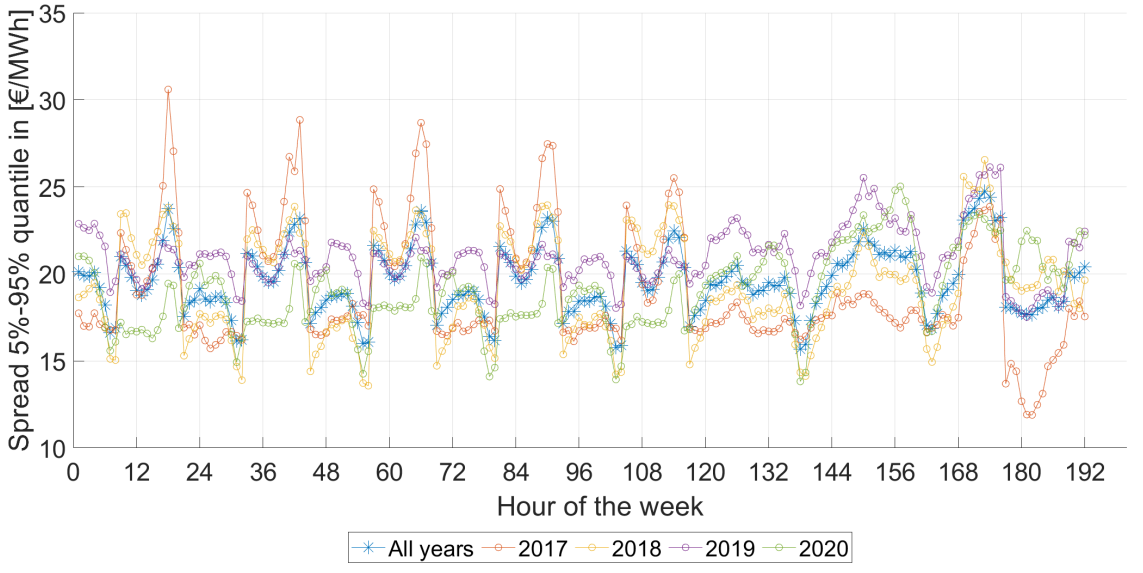


图9： 一周各时段预测的5%分位数与95%分位数之间的平均价差 .

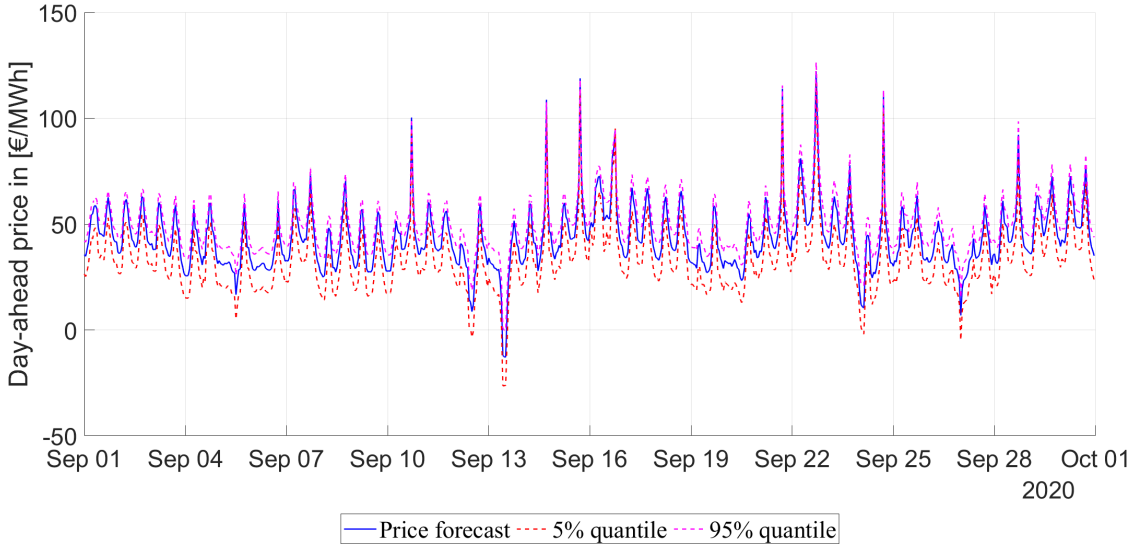


图10： 日前电价预测、下界预测（5%分位数）和上界预测（95%分位数）。

图 11 显示，星期日和公共假日的负价格发生概率最高，某些时段可达到 10% 至 15%。这反映出，在实际日前价格中，这些时段出现负价格的情况更为频繁。这是因为这些时段的电力需求相对较低，从而增加了负价格事件发生的可能性。星期一凌晨时段的负价格发生概率也相对较高（超过 5%），而一周其他日期凌晨时段的负价格发生概率则略高一些。这种模式反映了凌晨时段电力需求的低负荷特征，以及夜间和清晨通常较高的风电馈入，这些因素可能导致价格下降。在星期日和节假日，由于需求减少，白天光伏发电和持续的风电馈入可能增加负价格发生的可能性。

工作日（周一至周五）的概率值在水平和模式上相似，因为这些天的基本参数相近。根据密度预测确定概率后，可以制定竞价策略，

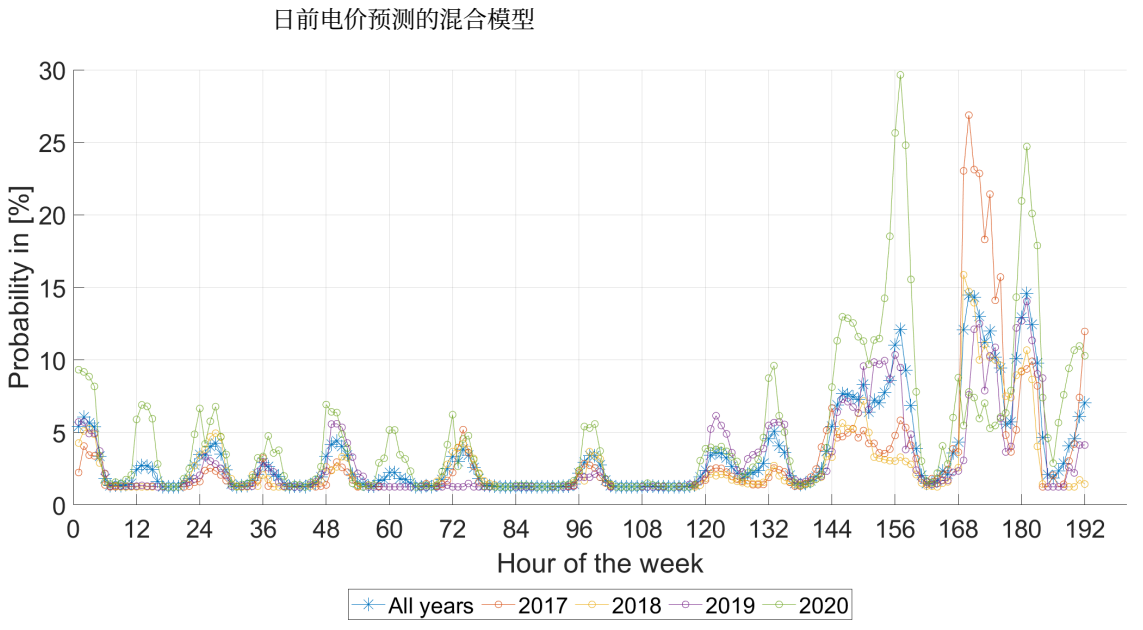


图 11：负日前价格的概率。

优化投资组合对冲和发电厂使用。负价格或大幅价差的风险可以通过概率陈述纳入策略。

总之，结果表明，该混合模型能够提供高质量的点预测和概率电价预测，同时还具备两项优于以往方法的额外优势。技术经济能源系统模型的整合有助于更深入地理解（能源）市场，而随机方法的应用则能够通过概率电价预测量化这些市场中的不确定性，并提供有关负价格及其他市场现象发生概率的宝贵信息，帮助市场参与者做出明智决策并降低风险。

6. 模型单个步骤的分析

本文提出的混合模型是多层次的。本节将介绍该混合模型各个组成部分（步骤）的一些情况。第 6.1节展示日前负荷预测的结果。第6.2节对能源系统优化步骤之后的价格估计器进行评估。最后，第 6.3节指出了六个独立价格估计器的误差预测所产生的影响。

6.1. 负荷预测

本节介绍负荷预测的结果，其中我们将输电系统运营商的负荷预测的均方根误差和平均绝对误差与表3中的改进负荷预测进行比较，结果显示均方根误差提高了 19%，平均绝对误差提高了 27%。有关更详细的分析，请参阅Möbius 等 [2023]。

6.2. 能源系统优化步骤后的价格估计器

在优化能源系统模型时，我们的混合模型根据需求约束的对偶变量生成价格估计器（参见第 4.3 节）。通过计算对偶变量，该模型优化能源行业复杂的技术经济相互依赖关系。除了发电成本外，该模型还考虑机组组合决策，例如发电机组的启动和停机、储能运行、与邻国之间电力输送的限制、供热需求以及调节电力的供应。所有这些技术经济相互依赖关系共同决定最终的电价，并且在我们的模型中至关重要，因为它们反映了市场上可观测到的高价格峰值以及低电价甚至负电价。

表 3原始输电系统运营商日前负荷预测（TSO）和改进的日前负荷预测（改进）的均方根误差和平均绝对误差，见 $[MWh]$。

	year	all	2016	2017	2018	2019	2020
RMSE	TSO	2,335.49	2,596.48	1,802.61	2,360.51	2,454.70	2,383.23
	改进	1,881.64	1,872.44	1,483.45	2,043.87	1,665.76	1,859.29
	改进百分比	19.43	27.89	17.71	13.41	32.14	21.98
MAE	TSO	1,797.05	2,028.44	1,396.45	1,726.67	1,951.00	1,881.84
	改进	1,320.83	1,302.17	1,106.12	1,372.55	1,283.17	1,405.71
	改进百分比	26.50	35.80	20.79	20.51	34.23	25.30

表4
能源系统优化步骤误差的描述性统计，单位为[欧分/兆瓦时]。 =

	所有年份	2016	2017	2018	2019	2020
mean	-2.17	-2.23	-3.06	0.02	-0.73	-4.84
中位数	-1.75	-2.24	-2.75	0.71	-0.23	-3.94
最小值	-143.85	-143.85	-102.45	-81.45	-80.01	-79.05
最大值	105.73	50.62	105.73	55.19	42.87	86.76
5%分位数	-15.30	-11.81	-17.77	-13.24	-12.78	-19.50
95%分位数	9.60	7.47	10.68	11.40	9.75	6.05
Std.	9.25	7.26	11.62	8.61	8.05	9.31
RMSE	9.50	7.60	12.01	8.61	8.09	10.49
MAE	6.00	4.83	7.09	5.91	5.07	7.13

表 4 展示了能源优化步骤后价格估计器的价格估计误差描述性统计，以及其均方根误差和平均绝对误差。对于所有年份，均方根误差为9。50 =欧分/兆瓦时，平均绝对误差为6。00 =欧分/兆瓦时。2016年可以观察到最低误差，而2017年可以观察到最高误差。与文献中的先进电价预测模型相比，误差度量更大，并且误差仍呈现结构性行为。在大多数年份，误差的均值和中位数均为负值，这意味着*em.power*调度计算出的电价高于市场上的观测价格。误差的5%分位数为-15。30 =欧分/兆瓦时，误差的95%分位数为9。60 =欧分/兆瓦时，90%的误差值落在此区间内。

图 12 展示了整个时期按一周内各小时划分后的误差结构。负值表示对应时段的价格平均高估，正值表示对应时段的价格平均低估。图中显示，周末的误差往往会增加，同时还能观察到明显的日变化模式（另见图 13）。

模型往往会在夜间低估价格，而在早晨和傍晚时段高估价格；从2016年至2019年，所有年份都呈现出这一稳定模式。尽管2020年的误差结构有所不同，这可能与新冠肺炎疫情有关，但模型按一天中各小时计算的均方根误差与往年相比并无显著差异。我们还观察到午间误差率上升，以及下午4点至6点之间的晚间误差峰值；这一现象在2020年的数据中更为明显。

不过请注意，我们在此评估的是技术经济能源系统模型的短期价格预测。由于该模型无法从历史数据中学习，我们预计它能够反映这些预测期限内的整体市场状况，但无法生成非常准确的价格估计器。因此，该模型对可预测的误差模式表现尚可，这表明数据后处理步骤具有很大的潜力。此步骤的影响将在下一章中介绍。

图 14 展示了价格预测误差与多个外生变量之间的相关性。显然，风力发电量与能源优化步骤产生的价格预测误差显著相关，并解释了总体预测误差波动性的 14 %。

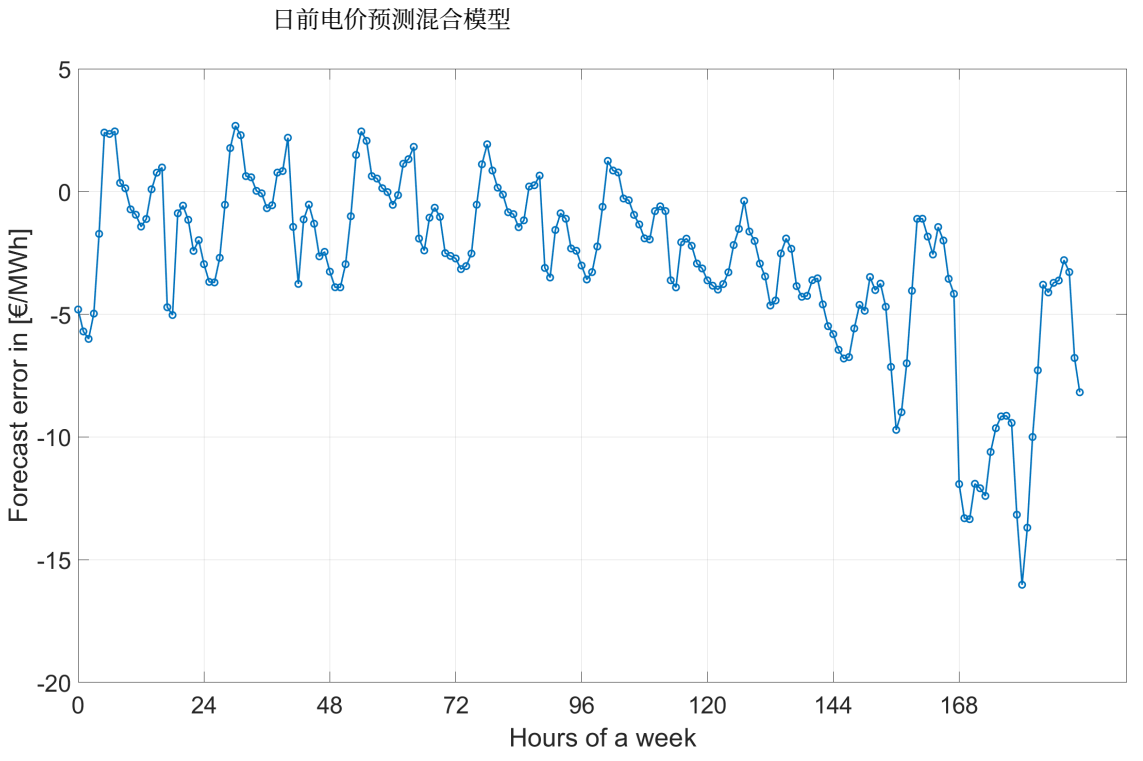


图12： [=能源系统优化步骤之后一周内各小时的平均价格估计误差（包括第168–192小时的公共假日），单位为欧分/兆瓦时]。

表 5混合模型在 [欧分/兆瓦时]下误差的描述性统计

=

	所有年份	2016	2017	2018	2019	2020
mean	-0.09	0.30	-0.15	1.12	-0.73	-0.99
中位数	-0.08	0.08	-0.23	1.06	-0.68	-0.50
最小值	-133.44	-133.44	-81.29	-54.41	-79.27	-69.43
最大值	86.10	43.76	74.30	52.72	35.48	86.10
5%分位数	-9.58	-6.64	-10.94	-9.09	-9.36	-11.51
95%分位数	10.07	8.48	11.21	11.70	9.08	8.24
Std.	7.38	5.82	8.79	7.20	7.01	7.56

6.3. 价格估计器的改进

在能源系统优化步骤之后价格估计器误差中检测到的系统性偏差和季节性模式，可以在后处理步骤中由随机模型捕获和预测。通过组合两类模型，2016–2020年的价格预测均方根误差降低了22 %。混合模型预测误差的统计数据表明，该误差近似呈对称分布，均值为-0.09 =欧分/兆瓦时，中位数为-0.08 =欧分/兆瓦时。后处理使价格预测误差趋近于零。同时，生成的最小误差和最大误差均逐年减少，并且在整个期间内也是如此，整体波动性也随之降低。因此，与能源系统优化步骤中的价格估计器相比，后处理模型使日前价格预测的标准差降低了20 %。

随机后处理步骤由六个单个子模型组成，如第4.4节所述。这些子模型中的每一个都会改进能源系统优化步骤后获得的价格估计器，从而显著提高预测质量，表 6 通过指定误差度量均方根误差和平均绝对误差对其进行了衡量。尽管在整个期间内，采用多变量框架的模型的误差度量低于采用单变量框架的模型，但在2019年和

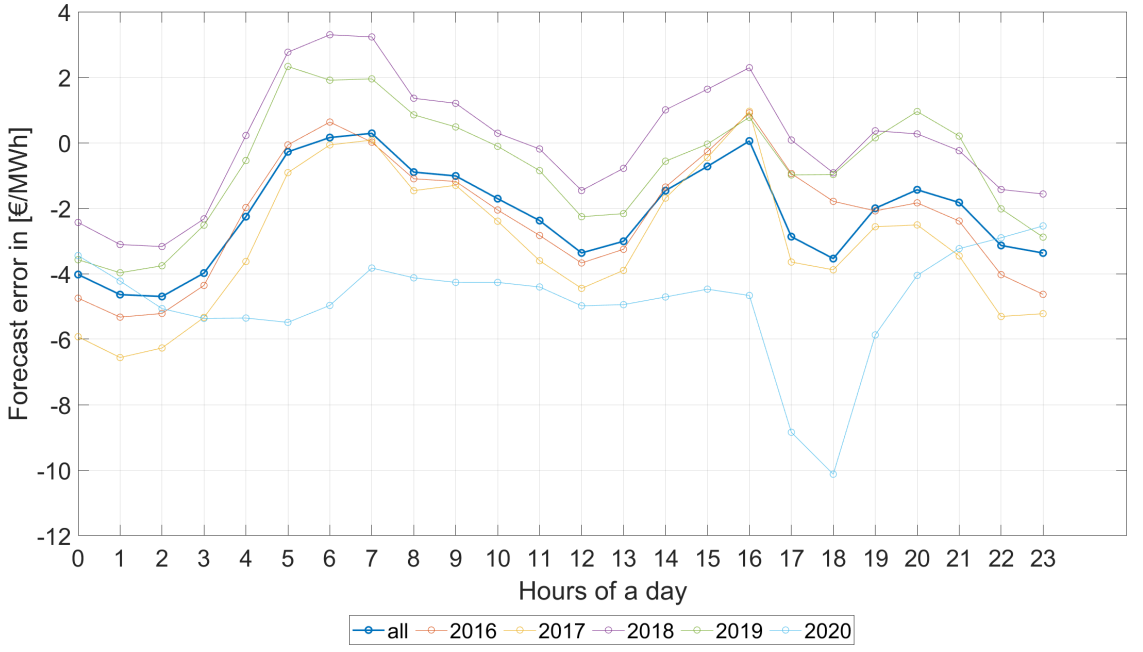


图13： [=C/MWh]下能源系统优化步骤之后，各年份一天中各小时的平均价格估计误差。

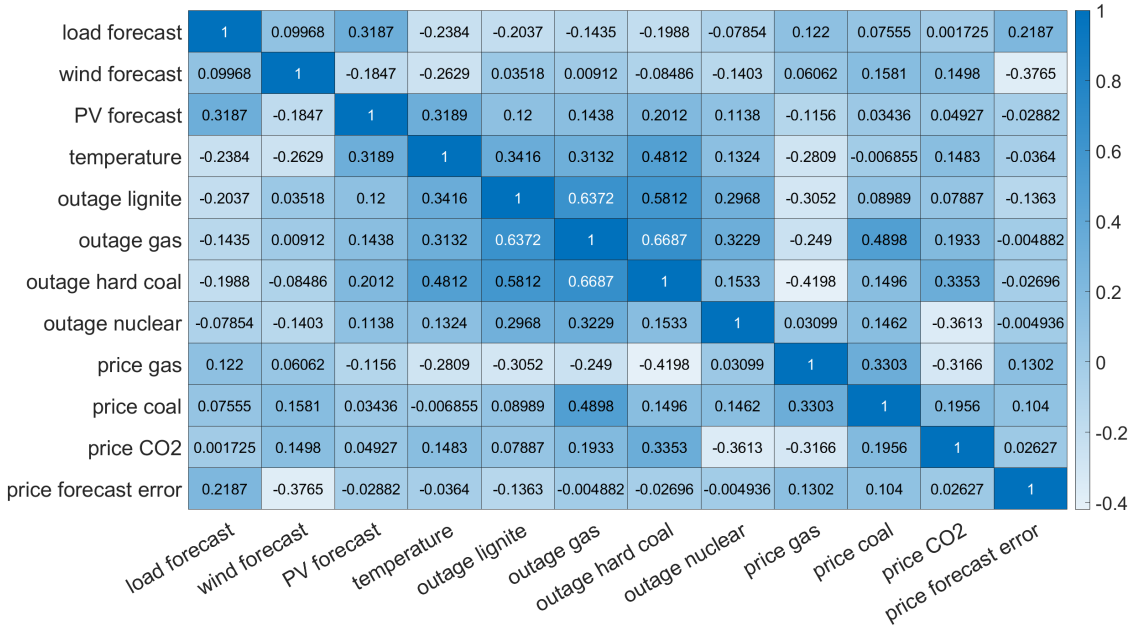


图14： 可能的外生变量与价格估计误差之间的相关性 .

2020年。图 15 展示了每个子模型的小时均方根误差。根据校准窗口的长度，单变量框架的模型在一天的前五个小时内优于多变量框架的模型。从这一时点起，多变量子模型取得了更低的误差度量。该结果与模型设定的预期一致。在单变量框架中，时间序列被视为高频数据集，因此预测包含截至最后一个可用小时的数据。然而，这也要求预测未来24个值，从而增加了不准确性。在多变量框架中，由于划分为

表6[欧分/兆瓦时]误差时间序列的均方根误差和平均绝对误差（误差度量）。
=

		初始	后处理							
		ESM	UV44w	UV48w	UV52w	MV44w	MV48w	MV52w	组合	
RMSE	所有年份	9.50	7.57	7.59	7.59	7.51	7.50	7.51	7.38	22 %
	2016	7.60	6.07	6.13	6.09	5.95	5.94	5.92	5.82	23 %
	2017	12.01	9.22	9.20	9.13	8.92	8.86	8.86	8.79	27 %
	2018	8.61	7.49	7.47	7.54	7.40	7.38	7.40	7.28	15 %
	2019	8.09	7.11	7.13	7.14	7.21	7.22	7.21	7.05	13 %
	2020	10.49	7.65	7.69	7.71	7.77	7.79	7.85	7.63	27 %
MAE	所有年份	6.00	4.73	4.75	4.75	4.75	4.74	4.74	4.60	23 %
	2016	4.83	3.69	3.71	3.69	3.64	3.62	3.61	3.48	28 %
	2017	7.09	5.55	5.55	5.51	5.37	5.31	5.28	5.25	26 %
	2018	5.91	5.20	5.18	5.21	5.24	5.23	5.25	5.07	14 %
	2019	5.07	4.46	4.49	4.49	4.61	4.61	4.59	4.43	13 %
	2020	7.13	4.77	4.80	4.83	4.91	4.94	4.99	4.77	33 %

24条独立时间序列，因为这种划分意味着即使是面向未来的短期步骤也需要进行预测。不过，最新的可用信息仅得到了部分考虑。

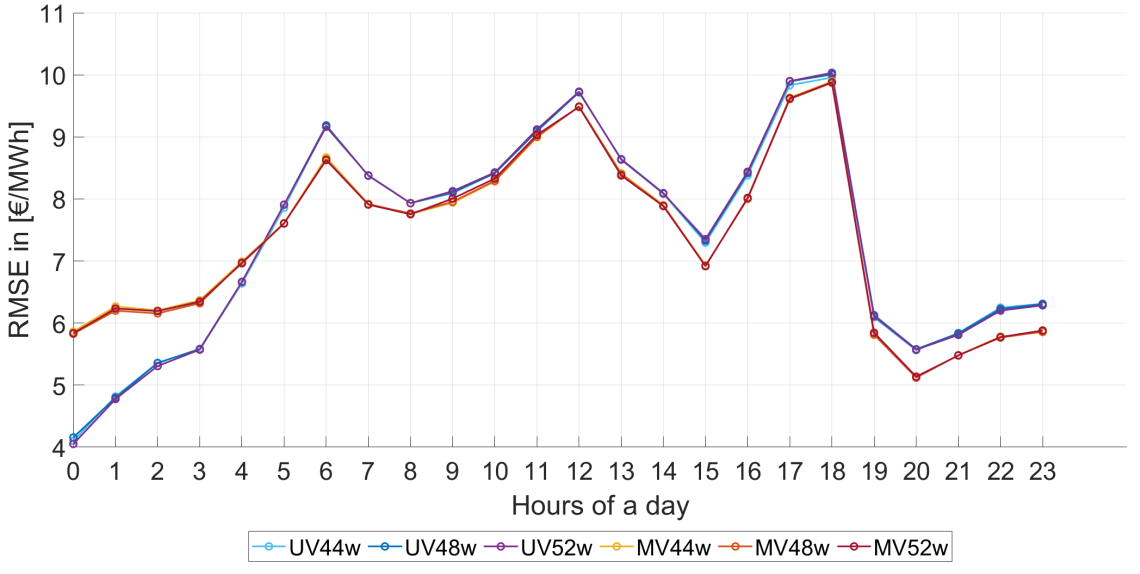


图15： [=欧分/兆瓦时]下，单变量和多变量后处理子模型在一天中各小时的均方根误差。

单变量和多变量框架各自的特征也解释了一天中各小时的平均误差（见图A.4）以及一周中各小时的平均误差（见图A.5）。多变量子模型的预测未显示出数值低估或高估的明显日模式，而单变量子模型则呈现出与能源系统优化步骤后价格估计器相似的日结构：凌晨4至8点和下午2至4点低估，而下午17至18点以及21点以后则存在高估。例如，按周内各小时统计的平均误差显示，公共假日的高估和低估幅度显著更大。但值得注意的是，这一趋势在单变量和多变量子模型的所有预测中均有所体现。

表6 呈现了由各个子模型的预测组合而成的价格预测的误差度量。因此，它描述了混合模型的最终价格预测。该价格预测的均方根误差和平均绝对误差低于各个价格预测在每个单独年份以及整个期间的误差度量。

使用*epftoolbox* 中的Lago等人 [2021] 提出的多变量 Diebold-Mariano 检验，通过假设检验比较不同模型的预测结果，并确定某一预测的准确性是否显著高于其他预测，我们证明了组合子模型显著优于全部六个单个子模型。检验结果如图16所示，热图展示了以下假设的 p 值：Y 轴模型的预测是否显著比 X 轴模型的预测更准确。接近

零表示 X 轴上的模型的预测准确度显著高于 Y 轴上的模型。检验表明，在预测日前现货价格时，后处理模型和每个随机子模型均显著优于经过随机预处理和交织处理修改的能源系统模型，也就是说，显著优于第三步之后的混合模型。

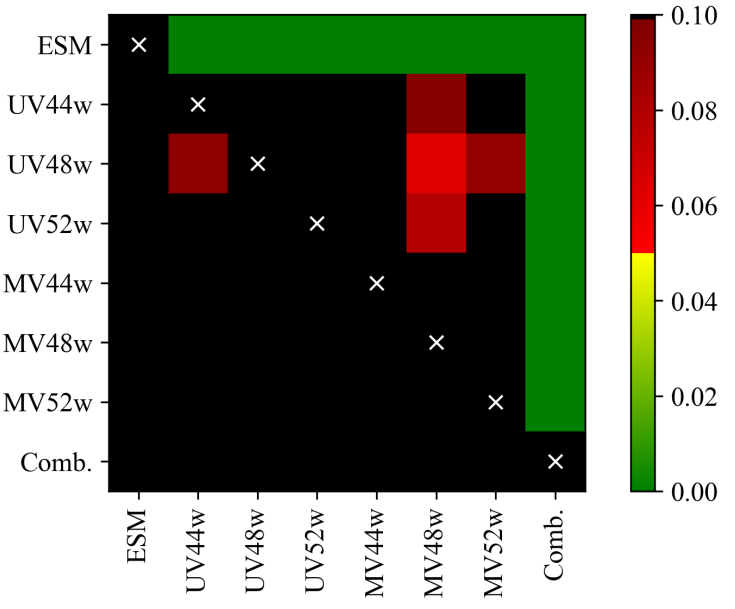


图16： 随机后处理步骤所有部分的 Diebold-Mariano 检验 p值（输出源自 *epftoolbox*，作者为 Lago等人 [2021]）。

为了更好地归因并理解随机后处理步骤的影响，我们还确定了一天中每个小时和一周中每天的均方根误差改进情况，如图17所示。白天和工作日结构在能源系统优化步骤之后的价格估计器误差中可以观察到，并且在改进情况中表现得十分明显。这在很大程度上是由于随机模型中的季节性成分以及受影响的自回归结构。值得注意的是，从百分比来看，最显著的改进出现在夜间。此外，周末，尤其是节假日，也取得了相对较高水平的改进。对于价格估计器而言，这些正是平均误差最显著的小时和日期，也就是说，这些时段具有较大的平均误差，因此也最具改进潜力。通过减少这一误差来提升日前电价预测，是结合技术经济能源系统建模与随机建模的关键目标。

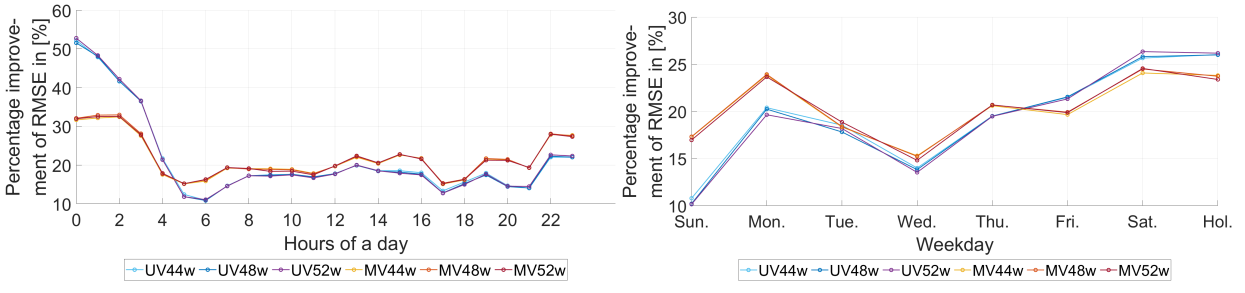


图 17： 一天中每个小时（左）和一周中每天（右）日前电价预测的平均均方根误差（RMSE）改善幅度。

7. 结论

本文介绍了一种用于短期电价预测的新型混合模型，该模型将随机建模与（基础性的）技术经济能源系统建模相结合。该模型由四个步骤组成。首先，通过随机数据预处理，日前负荷预测得到显著改进，并生成提前两天的负荷预测。其次，利用参数密度预测步骤，通过三个不同的可能情景，对提前两天的小时负荷消耗进行分位数预测，从而扩展该负荷预测。第三，将建模量和输入数据输入*em.power*调度，这是一个经过调整、用于预测日前价格估计值的技术经济市场模型。第四，通过随机数据后处理，利用随机模型降低能源优化步骤中价格估计器的误差，并以概率预测对其进行补充。

本文提出的混合模型结合了随机模型（优势：使用价格数据进行训练）和能源系统模型（优势：基于经济理论得出超越价格本身的洞见）的优点。该模型可以利用全球大多数能源市场的数据进行参数化，从而为各种能源系统提供洞见。我们使用2016年1月至2020年12月的德国日前电力市场数据验证了其性能。该模型的价格预测年均均方根误差为7.38 =欧分/兆瓦时，年均平均绝对误差为4.60 =欧分/兆瓦时。该混合模型的预测准确率超过了文献中的大多数基准模型，并与文献中确定的最佳统计基准模型相当。这证实，技术经济能源系统模型可以为电价预测提供有价值的输入，即使是在短期预测中也是如此。

我们对价格预测质量进行了深入分析，发现价格预测误差与价格水平之间存在显著关系。多年来，最高均方根误差始终出现在价格最低的20%和最高的20%对应的时段。这一结果可以通过启动成本及其对小时价格的影响来解释，正如第5节所述，并且自然会影响所有预测模型。

此外，我们的模型还提供校准密度预测以及点预测。例如，借助这些预测，可以量化价格为负的事件发生的概率。我们的分析发现，星期日和公共假日的某些时段出现负价格的概率超过 10 %。

未来研究应重点确定价格预测在不同提前期情形下最适合的模型类型。正如本文第一部分所述，技术经济模型在处理结构突变方面具有相对优势，因此更适合长期预测（例如数年后的预测）；而随机模型，无论是基于时间序列的模型还是基于人工智能的模型，都更适合短期预测（例如日内预测）。不过，本文针对日前预测提出的混合模型，以及其他结合上述两类模型优势的类似方法，如果经过适当调整，也可能适用于不同提前期下的精准预测。因此，我们认为，进一步探索和开发混合模型至关重要，以发挥其潜力，在各种提前期下提供精准预测，使其成为现有预测方法的有益补充。

数据可用性

本文所用的数据和源代码可通过以下链接公开获取：
https://github.com/ProKoMoProject/A-hybrid-model-for-day-ahead-electricity-price-forecasting-combining-fundamental-and-stochastic-mod。

致谢

本研究得到了德国联邦经济事务与气候行动部的支持，资助来自第六能源研究计划下系统分析研究网络中的 ProKoMo 项目。

参考文献

Rafał Weron。《电价预测：对最先进方法的综述及未来展望》。《国际预测杂志》，30(4)：1030–1081，2014年。ISSN 0169-2070。doi：10.1016/j.ijforecast.2014.08.008。

Jakub Nowotarski 和 Rafał Weron。电价预测的最新进展：概率预测综述。《可再生与可持续能源评论》，81：1548–1568，2018年1月。ISSN 18790690。doi：10.1016/j.rser.2017.05.234。Tao Hong、Pierre Pinson、Shu Fan、Hamidreza Zareipour、Alberto Troccoli 和 Rob J. Hyndman。2014年全球能源概率预测竞赛及其后续发展。《国际预测杂志》，32：896–913，2016年。ISSN 01692070。doi：10.1016/J.IJFORECAST.2016. 02.001。Tao Hong、Pierre Pinson、Yi Wang、Rafał Weron、Dazhi Yang 和 Hamidreza Zareipour。能源预测：综述与展望。IEEE开放获取电力与能源期刊，7：376–388，2020年。doi：10.1109/OAJPE.2020.3029979。N. Amjady 和 M. Hemmati。能源价格预测：此类预测面临的问题与建议。IEEE电力与能源杂志，4(2)：20–29，2006年。doi：10.1109/MPAE.2006.1597990。Sanjeev Kumar Aggarwal、Lalit Mohan Saini 和 Ashwani Kumar。放松管制市场中的电价预测：综述与评估。《国际电力与能源系统杂志》，31(1)：13–22，2009年。ISSN 0142-0615。doi：10.1016/j.ijepes.2008.09.003。Rafał Weron 和 Florian Ziel。电价预测。载于《劳特利奇能源经济学手册》，第506–521页。劳特利奇，2019年9月。doi：10.4324/9781315459653-36。Fotios Petropoulos、Daniele Apiletti、Vassilios Assimakopoulos、Mohamed Zied Babai、Devon K. Barrow、Souhaib Ben Taieb、Christoph Bergmeir、Ricardo J. Bessa、Jakub Bijak、John E. Boylan、Jethro Browell、Claudio Carnevale、Jennifer L. Castle、Pasquale Cirillo、Michael P. Clements、Clara Cordeiro、Fernando Luiz Cyrino Oliveira、Shari De Baets、Alexander Dokumentov、Joanne Ellison、Piotr Fiszeder、Philip Hans Franses、David T. Frazier、Michael Gilliland、M. Sinan Gönül、Paul Goodwin、Luigi Grossi、Yael Grushka-Cockayne、Mariangela Guidolin、Massimo Guidolin、Ulrich Gunter、Xiaojia Guo、Renato Guseo、Nigel Harvey、David F. Hendry、Ross Hollyman、Tim Januschowski、Jooyoung Jeon、Victor Richmond R. Jose、Yanfei Kang、Anne B. Koehler、Stephan Kolassa、Nikolaos Kourentzes、Sonia Leva、Feng Li、Konstantia Litsiou、Spyros Makridakis、Gael M. Martin、Andrew B. Martinez、Sheik Meeran、Theodore Modis、Konstantinos Nikolopoulos、Dilek Önkal、Alessia Paccagnini、Anastasios Panagiotelis、Ioannis Panapakidis、Jose M. Pavía、Manuela Pedio、Diego J. Pedregal、Pierre Pinson、Patrícia Ramos、David E. Rapach、J. James Reade、Bahman Rostami-Tabar、Michał Rubaszek、Georgios Sermpinis、Han Lin Shang、Evangelos Spiliotis、Aris A. Syntetos、Priyanga Dilini Talagala、Thiyanga S. Talagala、Len Tashman、Dimitrios Thomakos、Thordis Thorarinsdottir、Ezio Todini、Juan Ramón Trapero Arenas、Xiaoqian Wang、Robert L. Winkler、Alisa Yusupova 和 Florian Ziel。预测：理论与实践。《国际预测杂志》，38(3)：705–871，2022年。ISSN 0169-2070。doi：https://doi.org/10.1016/j.ijforecast.2021.11.001。URL https://www.sciencedirect.com/science/article/pii/S0169207021001758。Rick Steinert 和 Florian Ziel。使用期货进行短期至中期日前电价预测。《能源杂志》，40：105–127，2019年。ISSN 01956574。doi：10.5547/01956574.40.1.rste。Jakub Nowotarski 和 Rafał Weron。长期季节性成分在日前电价预测中的重要性。《能源经济学》，57：228–235，2016年6月。ISSN 01409883。doi：10.1016/J.ENECO.2016.05.009。Florian Ziel 和 Rafał Weron。具有高维结构的日前电价预测：单变量与多变量建模框架。《能源经济学》，70：396–420，2018年。ISSN 0140-9883。doi：10.1016/j.eneco.2017.12.016。T. M. Christensen、A. S. Hurn 和 K. A. Lindsay。电价尖峰预测。《国际预测杂志》，28 411，2012年4月400 01692070 10 1016 2011 02 019：–。ISSN。doi：。/j.ijforecast.。。Michael Eichler、Oliver Grothe、Hans Manner 和 Tuerk Dennis。澳大利亚电力市场尖峰发生短期预测模型：比较研究。《能源市场杂志》，7(1)：55–81，2014年3月。doi：10.21314/JEM.2014.104。Hans Manner、Dennis Türk 和 Michael Eichler。多变量电价尖峰建模与预测。《能源经济学》，60 265，2016年255 0140 9883 10 1016 2016 10 006：–。ISSN -。doi：。/j.eneco.。。R.C. Garcia、J. Contreras、M. van Akkeren 和 J.B.C. Garcia。用于预测日前电价的GARCH预测模型。《IEEE电力系统汇刊》，20(2)：867–874，2005年。doi：10.1109/TPWRS.2005.846044。Emily Hickey、David G. Loomis 和 Hassan Mohammadi。使用ARMAX–GARCH模型预测小时电价：MISO枢纽的应用。《能源经济学》，34(1)：307–315，2012年。ISSN 0140-9883。doi：10.1016/j.eneco.2011.11.011。Silvano Bordignon、Derek W. Bunn、Francesco Lisi 和 Fany Nan。英国电价日前预测的组合。《能源经济学》，35：88–103，2013年1月。ISSN 01409883。doi：10.1016/j.eneco.2011.12.001。Peter Kosater 和 Karl Mosler。马尔可夫制度转换模型能否改善电力价格预测？来自德国日度电价的证据。《应用能源》，83(9)：943–958，2006年。ISSN 0306-2619。doi：10.1016/j.apenergy.2005.10.007。Michael Bierbrauer、Stefan Trück 和 Rafał Weron。使用制度转换模型对电价进行建模。《计算机科学讲义》，3039：859–867，2004年。ISSN 16113349。doi：10.1007/978-3-540-25944-2_111。Carlo Mari 和 Emiliano Mari。基于深度学习的能源大宗商品价格制度转换模型。《能源系统》，2022年。doi：10.1007/s12667-022-00515-6。Bartosz Uniejewski、Jakub Nowotarski 和 Rafał Weron。日前电价预测的自动变量选择与收缩。《能源》，9(8)，2016年。ISSN 1996-1073。doi：10.3390/en9080621。Jesus Lago、Grzegorz Marcjasz、Bart De Schutter 和 Rafał Weron。日前电价预测：最先进算法、最佳实践和开放获取基准综述。《应用能源》，293：116983，2021年。ISSN 0306-2619。doi：https://doi.org/10.1016/j.apenergy. 2021.116983。URL https://www.sciencedirect.com/science/article/pii/S0306261921004529。Ioannis P. Panapakidis 和 Athanasios S. Dagoumas。应用基于人工神经网络的模型进行日前电价预测。《应用能源》，172：132–151，2016年6月。ISSN 03062619。doi：10.1016/j.apenergy.2016.03.089。Jesus Lago、Fjo De Ridder、Peter Vrancx 和 Bart De Schutter。欧洲日前电价预测：考虑市场一体化的重要性。《应用能源》，211：890–903，2018年2月。ISSN 03062619。doi：10.1016/j.apenergy.2017.11.098。Nima Amjady。通过一种新的模糊神经网络预测电力市场日前电价。《IEEE电力系统汇刊》，21(2)：887–896，2006年。doi：10.1109/TPWRS.2006.873409。Grzegorz Marcjasz、Bartosz Uniejewski 和 Rafał Weron。使用NARX神经网络进行日前电价预测时长期季节性成分的重要性。《国际预测杂志》，35(4)：1520–1532，2019年。ISSN 0169-2070。doi：

10.1016/j.ijforecast.2017.11.009。Malte Lehna、Fabian Scheller 和 Helmut Herwartz。日前电价预测：考虑外生回归变量的时间序列模型与神经网络模型比较。《能源经济学》，106:105742，2022年。ISSN 0140-9883。doi：10.1016/j.eneco.2021.105742。Ramiz Qussous、Nick Harder 和 Anke Weidlich。通过反映市场相互关系和面向灵活性的竞价策略理解电力市场动态。《能源》，15(2)，2022年。ISSN 1996-1073。doi：10.3390/en15020494。Felix Muesgens。使用动态多区域调度模型量化德国电力批发市场的市场势力。《The Journal of Industrial Economics》，54(4):471–498，2006年。doi：10.1111/j.1467-6451.2006.00297.x。Severin Borenstein、James B. Bushnell 和 Frank A. Wolak。衡量加利福尼亚州重组批发电力市场的市场低效。《American Economic Review》，92(5):1376–1405，2002年12月。doi：10.1257/000282802762024557。Dogan Keles、Massimo Genoese、Dominik Möst、Sebastian Ortlieb 和 Wolf Fichtner。风电馈入与电力现货价格的综合建模方法。《Energy Policy》，59:213–225，2013年。ISSN 0301-4215。doi：https://doi.org/10.1016/j.enpol.2013.03.028。URL https://www.sciencedirect.com/science/article/pii/S0301421513001821。Lion Hirth。可变可再生能源的市场价值：太阳能和风电波动性对其相对价格的影响。《能源经济学》，38:218–236，2013年。ISSN 0140-9883。doi：10.1016/j.eneco.2013.02.004。Frank Sensfuß、Mario Ragwitz 和 Massimo Genoese。优先排序效应：可再生能源发电对德国现货市场价格影响的详细分析。《Energy Policy》，36(8):3086–3094，2008年。ISSN 0301-4215。doi：10.1016/j.enpol.2008.03.035。Christian Pape、Si mon Hagemann 和 Christoph Weber。基本面是否足够？解释德国日前与日内电力市场的价格变化。《能源经济学》，54:376–387，2016年2月。ISSN 0140-9883。doi：10.1016/J.ENECO.2015.12.013。Felix Muesgens。具有二氧化碳排放交易和高比例可再生能源的电力系统中的均衡价格与投资。《能源经济学》，86:104107，2020年。ISSN 0140-9883。doi：10.1016/j.eneco.2018.07.028。Richard Green 和 Nicholas Vasilakos。大量间歇性发电条件下的市场行为。《Energy Policy》，38(7):3211–3220，2010年。ISSN 0301-4215。doi：10.1016/j.enpol.2009.07.038。具有《Regular Papers》的电力市场大规模风电。Alan D. Lamont。评估间歇性发电技术的长期系统价值。《能源经济学》，30:1231，2008年3月:1208，0140 9883，10 1016 2007 02 007（）：–。ISSN -。doi：./j.eneco.。。Steffi Misconel、Christoph Zöphel 和 Dominik Möst。评估需求响应在脱碳能源系统中的价值——大规模模型应用。《应用能源》，299:117326，2021年。ISSN 0306-2619。doi：10.1016/j.apenergy.2021.117326。Dana Kirchem、Muireann Á. Lynch、Valentin Bertsch 和 Eoin Casey。使用过程模型和能源系统模型对需求响应进行建模：废水处理在能源—水纽带中的潜在应用。《应用能源》，260:114321，2020年。ISSN 0306-2619。doi：10.1016/j.apenergy.2019.114321。Adriaan Hendrik van der Weijde 和 Benjamin F. Hobbs。规划输电以接纳可再生能源的经济学：使用两阶段优化评估灵活性及忽视不确定性的成本。《能源经济学》，34(6):2089–2101，2012年。ISSN 0140-9883。doi：10.1016/j.eneco.2012.02.015。Fabian Scheller 和 Thomas Bruckner。市级能源系统优化：建模方法与挑战分析。《可再生与可持续能源评论》，105:444–461，2019年。ISSN 1364-0321。doi：10.1016/j.rser.2019.02.005。Muireann Lynch、Mel T. Devine 和 Valentin Bertsch。电转气在未来能源系统中的作用：市场与组合效应。《能源》，185:1197–1209，2019 年。ISSN 0360-5442。doi：10.1016/j.energy.2019.07.089。Smaranda Sgarciu、Daniel Scholz 和 Felix Muesgens。二氧化碳价格如何加速脱碳——德国燃煤发电案例。《Energy Policy》，173:113375，2023年。ISSN 0301-4215。doi：10.1016/j.enpol.2022.113375。Leonie Sara Plaga 和 Valentin Bertsch。能源系统模型中评估气候不确定性的方法——系统性文献综述。《应用能源》，331:120384，2023年。ISSN 0306-2619。doi：10.1016/j.apenergy.2022.120384。Mariano Ventosa、Álvaro Baíllo、Andrés Ramos 和 Michel Rivier。电力市场建模趋势。《Energy Policy》，33(7):897–913，2005年。ISSN 0301-4215。doi：10.1016/j.enpol.2003.10.013。Abhinav Aggarwal 和 M M Tripathi。一种新型混合方法：结合小波变换、时间序列时间延迟神经网络和误差预测算法进行日前电价预测。载于《2017年第6届电气工程计算机应用——最新进展国际会议（CERA）》，第199–204页，2017年。doi：10.1109/CERA.2017.8343326。Zihan Chang、Yang Zhang 和 Wenbo Chen。基于 Adam 优化的 LSTM神经网络与小波变换混合模型的电价预测。《能源》，187:115804，2019年。ISSN 0360-5442。doi：10.1016/j.energy.2019.07.134。Hangyang Cheng、Xiangwu Ding、Wuneng Zhou 和 Renqiang Ding。采用贝叶斯优化的德国能源交易所混合电价预测模型。《国际电力与能源系统杂志》，110:653–666，2019年。ISSN 0142-0615。doi：10.1016/j.ijepes.2019.03.056。Mehrdad Setayesh Nazar、Ashkan Eslami Fard、Alireza Heidari、Miadreza Shafie-khah 和 João P.S. Catalão。使用三阶段算法同时进行负荷与电价预测的混合模型。《Electric Power Systems Research》，165:214–228，2018年。ISSN 0378-7796。doi：10.1016/j.epsr.2018.09.004。Zhang Yang、Li Ce 和 Li Lian。结合小波变换、ARMA和基于核的极限学习机方法的混合电价预测。《应用能源》，190:291–305，2017年。ISSN 0306-2619。doi：10.1016/j.apenergy.2016.12.130。Javad Olamaee、Mohsen Mohammadi、Alireza Noruzi 和 Seyed Mohammad Hassan Hosseini。基于混合预测模型的日前电价预测。《Complexity》，21(S2):156–164，2016年。doi：10.1002/cplx.21792。Jinliang Zhang、Zhongfu Tan 和 Yiming Wei。短期电价预测的自适应混合模型。《应用能源》，258:114087，2020年。ISSN 0306-2619。doi：10.1016/j.apenergy.2019.114087。Rodrigo A. de Marcos、Antonio Bello 和 Javier Reneses。短期电价预测：基本面建模与计量经济建模的混合。《Electric Power Systems Research》，167:240–251，2019年。ISSN 0378-7796。doi：10.1016/j.epsr.2018.10.034。Virginia Gonzalez、Javier Contreras 和 Derek W. Bunn。使用基本面—计量经济模型的电价预测。《IEEE电力系统汇刊》，27(1):363–372，2012年。doi：10.1109/TPWRS.2011.2167689。

Thomas Möbius、Mira Watermeyer、Oliver Grothe 和 Felix Muesgens。利用更好的负荷预测增强能源系统模型。arXiv，2023年。doi：10.48550/ARXIV.2302.11017。Katarzyna Maciejowska 和 Jakub Nowotarski。用于GEFCom2014概率电价预测的混合模型。《国际预测杂志》，32(3):1051–1056，2016年。ISSN 0169-2070。doi：10.1016/j.ijforecast.2015.11.008。Rafał Weron 和 Adam Misiorek。现货电价预测：参数与半参数时间序列模型的比较。《国际预测杂志》，24(4):744–763，2008年。ISSN 0169-2070。doi：10.1016/j.ijforecast.2008.08.004。能源预测。Xia Chen、Zhao Yang Dong、Ke Meng、Yan Xu、Kit Po Wong 和 H. W. Ngan。利用极限学习机和自助法进行电价预测。《IEEE电力系统汇刊》，27(4):2055–2062，2012年。doi：10.1109/TPWRS.2012.2190627。Can Wan、Zhao Xu、Yelei Wang、Zhao Yang Dong 和 Kit Po Wong。电价概率预测的混合方法。《IEEE智能电网汇刊》，5(1):463–470，2014年。doi：10.1109/TSG.2013.2274465。Mehdi Rafiei、Taher Niknam 和 Mohammad Khooban。通过改进的克隆选择算法和小波预处理进行概率电价预测。《神经计算与应用》，28，2017年12月。doi：10.1007/s00521-016-2279-7。Abbas Khosravi、Saeid Nahavandi 和 Doug Creighton。量化基于神经网络的电价预测的不确定性。《应用能源》，112:120–129，2013年。ISSN 0306-2619。doi：10.1016/j.apenergy.2013.05.075。Jun Hua Zhao、Zhao Yang Dong、Zhao Xu 和 Kit Po Wong。电价区间预测的统计方法。《IEEE电力系统汇刊》，23(2):267–276，2008年。doi：10.1109/TPWRS.2008.919309。Ming Zhou、Z. Yan、Y.X. Ni、Gengyin Li 和 Y. Nie。通过扩展ARIMA方法及置信区间估计进行电价预测。《发电、输电与配电》，IEE Proceedings-，第187–195页，2006年4月。doi：10.1049/ip-gtd:20045131。Anastasios Panagiotelis 和 Michael Smith。使用多元偏斜t分布对日内电价进行贝叶斯密度预测。《国际预测杂志》，24:710–727，2008年10月。doi：10.1016/j.ijforecast.2008.08.009。Hans Manner、Farzad Alavi Fard、Armin Pourkhanali 和 Laleh Tafakori。使用动态藤 copula预测澳大利亚电价的联合分布。《能源经济学》，78:143–164，2019年。ISSN 0140-9883。doi：https://doi.org/10.1016/j.eneco.2018.10.034。Oliver Grothe、Fabian Kächele 和 Fabian Krüger。从点预测到多元概率预测：Schaake洗牌法在日前电价预测中的应用。《能源经济学》，第106602页，2023年。ISSN 0140-9883。doi：10.1016/j.eneco.2023.106602。Jakub Nowotarski 和 Rafał Weron。使用分位数回归和预测平均计算电力现货价格预测区间。《计算统计学》，30:791–803，2015年9月。ISSN 16139658。doi：10.1007/S00180-014-0523-0/FIGURES/3。Adam Misiorek、Stefan Trueck 和 Rafal Weron。电力现货价格的点预测与区间预测：线性与非线性时间序列模型。《非线性动力学与计量经济学研究》，10(3)，2006年。doi：10.2202/1558-3708.1362。Grzegorz Dudek。用于GEFCom2014概率电价预测的多层感知机。《国际预测杂志》，32(3):1057–1060，2016年。ISSN 0169-2070。doi：10.1016/j.ijforecast.2015.11.009。Katarzyna Maciejowska、Jakub Nowotarski 和 Rafał Weron。使用因子分位数回归平均进行电力现货价格概率预测。《国际预测杂志》，32(3):957–965，2016年。ISSN 0169-2070。doi：10.1016/j.ijforecast.2014.12.004。Rafał Weron。电力负荷和电价建模与预测：统计方法。Wiley金融丛书。Wiley父子出版公司，奇切斯特

[ ], 等。2006 。 047005753 9780470057537
ISBN X；。Jakub Nowotarski 和 Rafał Weron。将分位数回归与预测平均相结合，以获得更准确的Nord Pool现货价格区间预测。载于《第11届欧洲能源市场国际会议》（EEM14），第1–5页，2014年。doi：10.1109/EEM.2014.6861285。Grzegorz Marcjasz、Bartosz Uniejewski 和 Rafał Weron。使用NARX网络进行概率电价预测：结合点预测还是概率预测？《国际预测杂志》，36(2):466–479，2020年。ISSN 0169-2070。doi：10.1016/j.ijforecast.2019.07.002。Bartosz Uniejewski、Grzegorz Marcjasz 和 Rafał Weron。日前电价预测中长期季节性成分的重要性：第二部分——概率预测。《能源经济学》，79:171–182，2019年。ISSN 0140-9883。doi：10.1016/j.eneco.2018.02.007。变化环境中的能源市场动态。Bartosz Uniejewski 和 Rafał Weron。用于概率电价预测的正则化分位数回归平均。《能源经济学》，95:105121，2021年。doi：10.1016/j.eneco.2021.105121。Sandbag。二氧化碳排放配额，2020年。URL https://sandbag.be/index.php/carbon-price-viewer/。访问日期：03-01-2022。Regelleistung.net。招标容量清单，2018年。URL https://www.regelleistung.net/ext/。访问日期：12-12-2021。Andreas Schröder、Friedrich Kunz、Jan Meiss、Roman Mendelevitch 和 Christian von Hirschhausen。到2050年的当前及预期发电成本。《DIW数据文档》，68，2013年。ISSN 1861-1532。开放电力系统数据。天气数据包。2020-09-16版，2020年。URL https://doi.org/10.25832/weather_data/ 2020-09-16。访问日期：15-05-2020。ENTSO-E透明度平台。总负荷——日前/实际，2021年。URL https://transparency.entsoe.eu/。访问日期：20-12-2021。德国联邦统计局。工业产品出厂价格（国内销售）。轻质燃料油、汽油和柴油价格，2021年。URL https://www.destatis.de/DE/Themen/Wirtschaft/Preise/Erzeugerpreisindex-gewerbliche-Produkte/_inhalt.html。访问日期：25-01-2021。ENTSO-E。TYNDP 2018情景报告，2018年。URL https://tyndp.entsoe.eu/tyndp2018/scenario-report。访问日期：23-02-2023。EEX。欧洲能源交易所：历史天然气价格数据，2021年。访问日期：03-01-2022。德国联邦网络局。联邦网络局电厂名单，2021年。URL https://www.bundesnetzagentur.de/。访问日期：03-01-2022。德国联邦环境局。德国电厂数据库，2020年。URL https://www.umweltbundesamt.de/dokument/datenbank-kraftwerke-in-deutschland。访问日期：03-01-2022。EBC。Europe Beyond Coal：欧洲燃煤电厂数据库，2021年1月25日，2021年。URL https://beyond-coal.eu/database/。访问日期：25-01-2021。ENTSO-E透明度平台。按发电类型划分的装机容量，2021年。URL https://transparency.entsoe.eu/。访问日期：20-12-2021。

开放电力系统数据。国家发电容量数据包。版本 2019-12-02，2020b。网址：https://doi.org/10.25832/ national_generation_ capacity/2019-12-02。访问日期：15-05-2020。欧盟委员会。欧盟统计局统计数据库，2021。网址：https://ec.europa.eu/eurostat/data/database。访问日期：12-12- 2021。ENTSO-E 透明度平台。按生产类型划分的实际发电量，2021c。网址：https://transparency.entsoe.eu/。访问日期：20-12-2021。ENTSO-E 透明度平台。日前输电容量预测，2021d。网址：https://transparency.entsoe.eu/。访问日期：20-12-2021。JAO联合分配办公室。影子拍卖的可用输电容量，2021。网址：https://www.jao.eu/implict-allocation。访问日期：20-12-2021。ENTSO-E 透明度平台。生产和发电机组不可用情况，2021年。网址：https://transparency.entsoe.eu/。访问日期：20-12-2021。ENTSO-E 透明度平台。日前发电预测，2021f。网址：https://transparency.entsoe.eu/。访问日期：20-12- 2021。ENTSO-E 透明度平台。日前价格，2021g。网址：https://transparency.entsoe.eu/。访问日期：20-12-2021。Mark Hellwig。参数化标准负荷曲线的发展与应用。博士论文，慕尼黑工业大学，2003年。博士论文，慕尼黑工业大学。德国能源署。德国能源署电网研究 II：2015—2020年可再生能源融入德国电力供应，并展望2025年，2010年。网址：https://www.dena.de/fileadmin/user_upload/Download/Dokumente/Studien___Umfragen/Endbericht_ dena-Netzstudie_II.PDF。Friedrich Kunz、Mario Kendziorski、Wolf-Peter Schill、Jens Weibezahn、Jan Zepter、Christian R. von Hirschhausen、Philipp Hauser、Matthias Zech、Dominik Möst、Sina Heidari、Björn Felten 和 Christoph Weber。用于德国系统建模的电力、供热和燃气部门数据。DIW数据文档 92，2017年。网址：http://hdl.handle.net/10419/173388。Katarzyna Maciejowska、Weronika Nitka 和 Tomasz Weron。通过增强负荷、风电和太阳能发电数据改进日前电价预测。《能源经济学》，99:105273，2021年7月。ISSN 01409883。doi：10.1016/j.eneco.2021.105273。Helmut Lütkepohl。《多元时间序列分析新论》。施普林格，柏林，2005年。ISBN 3540401725；3540262393；9783540262398；9783540401728。Rob J. Hyndman 和 George Athanasopoulos。《预测：原理与实践》。Otexts：澳大利亚墨尔本、肯塔基州列克星敦，2021年。ISBN 0987507133。访问日期：04-02-2022。George E. P. Box、Gwilym M. Jenkins、Gregory C. Reinsel 和 Greta M. Ljung。《时间序列分析：预测与控制》。威利概率与统计丛书。约翰·威立父子公司，新泽西州霍博肯，第五版，2015年。ISBN 9781118675021。doi：10.1111/jtsa.12194。Roger Koenker 和 Gilbert Bassett。回归分位数。《计量经济学》，46(1):33，1978年1月。doi：10.2307/1913643。Oliver Grothe。高阶相关无迹卡尔曼滤波器。《应用数学与计算》，219(17):9033–9042，2013年。ISSN 0096-3003，1873-5649。doi：10.1016/j.amc.2013.03.019。Felix Muesgens。利用动态多区域调度模型量化德国电力批发市场的市场势力。《产业经济学杂志》，54(4):471–498，2006b。doi：10.1111/j.1467-6451.2006.00297.x。Felix Muesgens 和 Karsten Neuhoff。电力市场动态约束建模及不确定风电出力的成本，2006年。Grzegorz Marcjasz、Tomasz Serafin 和 Rafał Weron。日前电价预测校准窗口的选择。《能源》，11:2364，2018年9月。ISSN 1996-1073。doi：10.3390/EN11092364。Victor DeMiguel、Lorenzo Garlappi 和 Raman Uppal。最优分散投资与朴素分散投资：1/n 投资组合策略的效率有多低？《金融研究评论》，22(5):1915–1953，2009年。doi：10.1093/rfs/hhm075。Ludwig Kuntz 和 Felix Muesgens。电力市场中多种技术启动成本的建模。《运筹学数学方法》，66(1):21–32，2007年。ISSN 14322994。doi：10.1007/s00186-007-0148-y。Tilmann Gneiting、Fadoua Balabdaoui 和 Adrian E Raftery。概率预测、校准与锐度。《英国皇家统计学会志：B辑（统计方法论）》，69(2):243–268，2007年。doi：10.1111/j.1467-9868.2007.00587.x。

A. 补充分析材料

本附录为以下章节提供补充图表5 和 6。

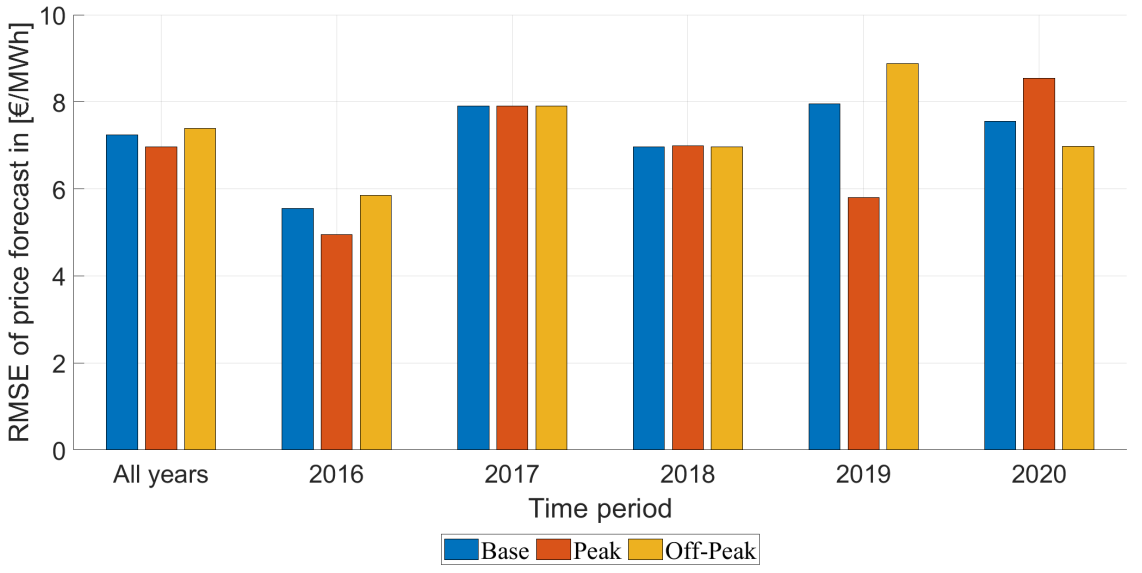


图 A.1： LEAR基准模型对基准时段、高峰时段和非高峰时段的日前电价预测的均方根误差。

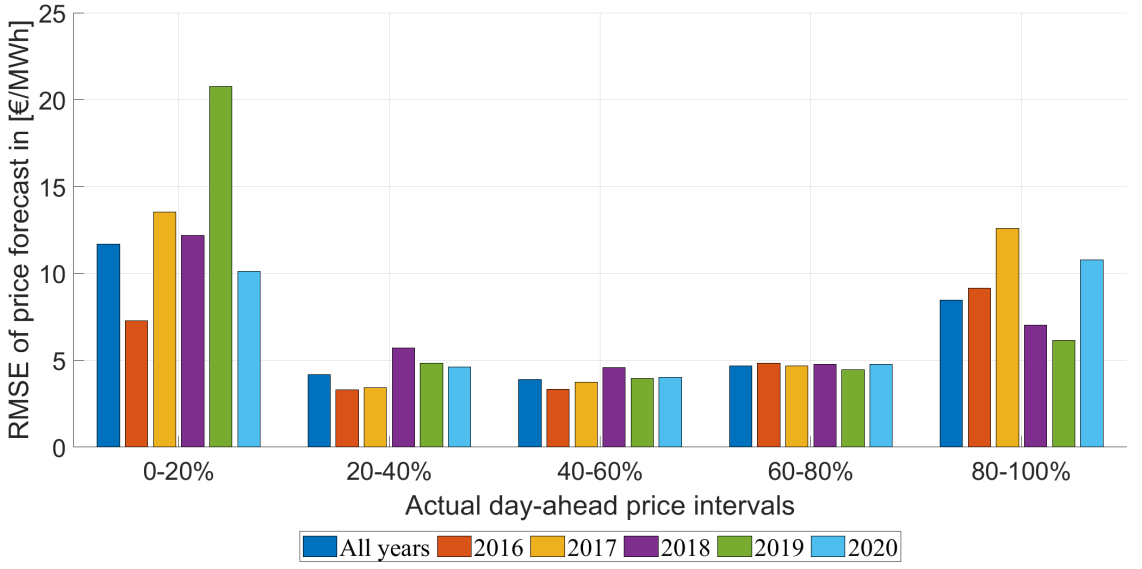


图 A.2： LEAR基准模型对处于不同日前价格分位数的时段进行日前电价预测的均方根误差。

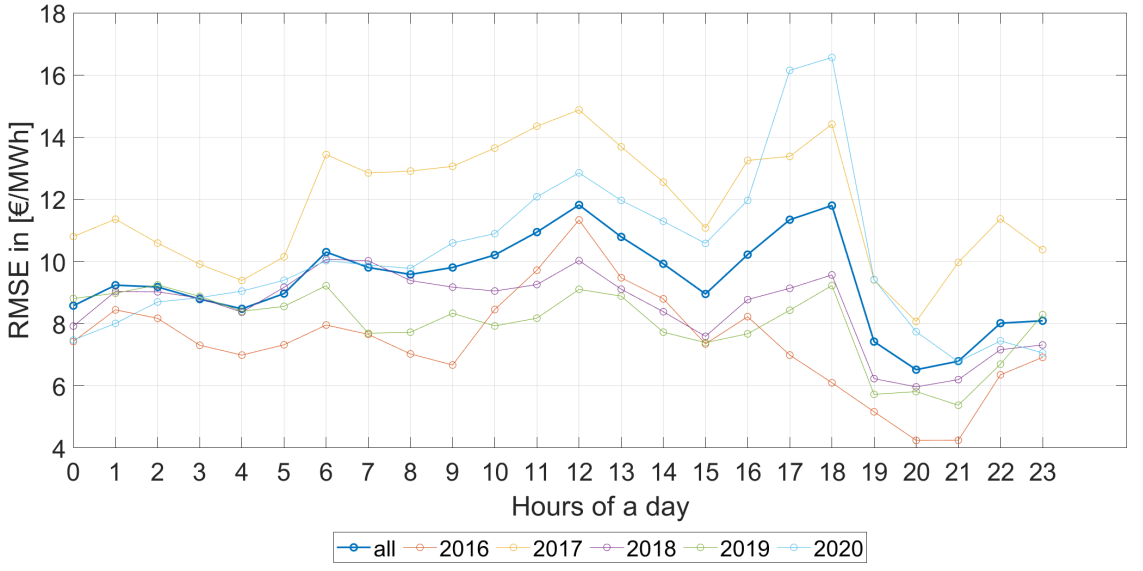


图 A.3： [=欧分/兆瓦时]能源系统优化步骤后，各年份每日每小时价格估计器的均方根误差。

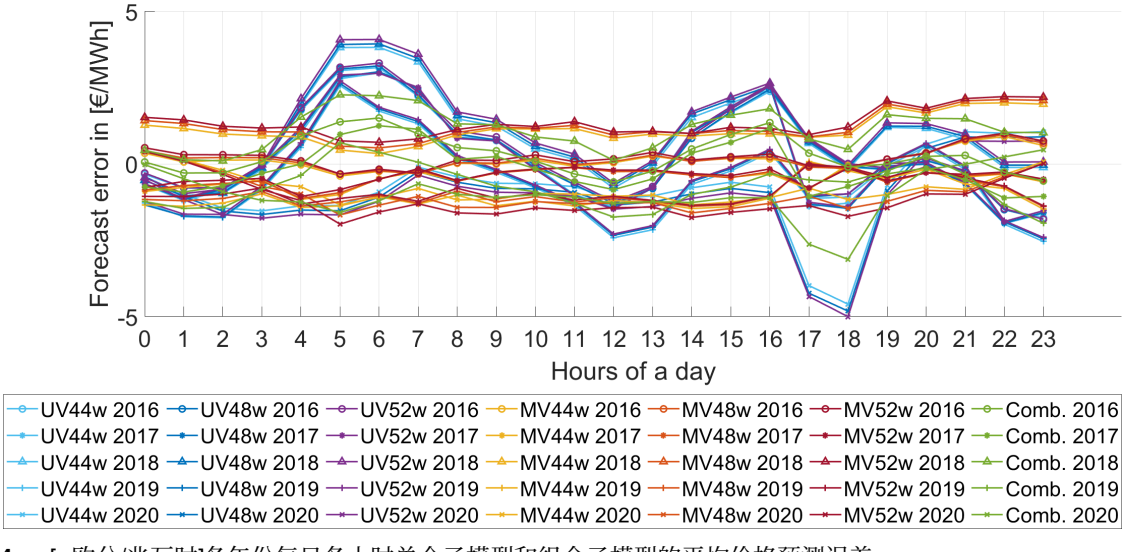


图 A.4： [=欧分/兆瓦时]各年份每日各小时单个子模型和组合子模型的平均价格预测误差。

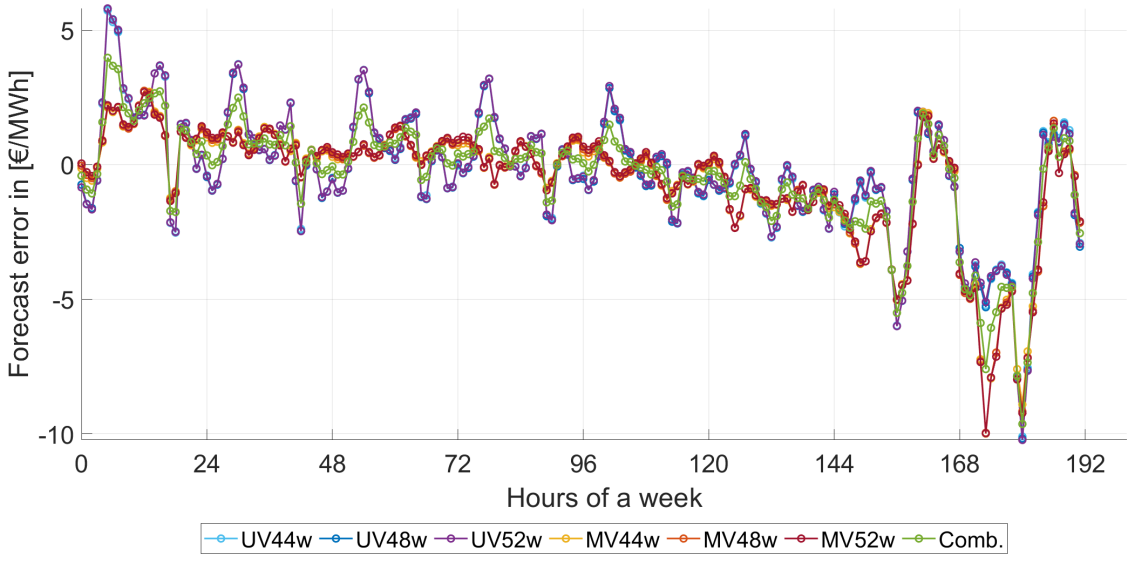


图 A.5： [=欧分/兆瓦时]的某一周各小时（包括节假日，第 168 至 192 小时）单个子模型和组合子模型的平均价格预测误差。

B. 能源系统模型符号表

Sets and indices

BP	Time blocks for primary control power
BS	Time blocks for secondary control power
D	Amount of days of a rolling window that includes $d+0, d+1, d+2$
$hfirst(H)$	First hour of a day
$hlast(H)$	Last hour of a day
$RES(I)$	Amount of intermittent renewables [Subset of I]
$STL(I)$	Amount of long-term storage [Subset of I]
$STM(I)$	Amount of mid-term storage [Subset of I]
H	Amount of hours of a day
I	Amount of electricity generation capacity clusters
$N, HR(I)$	Amount of hydro reservoirs [Subset of I]
N, NN	Amount of nodes

Parameters

η_i	Efficiency rate of a generation technology
$af_{i,n,d,h}$	Availability factor for generation capacities
$cap_{i,n,d,h}$	Installed generation capacity $[MW_{el}]$
cer	capacity-to-energy ratio for storage that operate actively within a three-day cycle $[MWh_{el}/MW_{el}]$
$chp_{i,n,d,h}$	Minimum electricity output of combined heat power units $[MWh_{el}/h]$
$curtc$	Costs for RES curtailment $[€/MWh_{el}]$
$g^{min}_{i,n,d,h}$	Minimum generation of a running unit
$ntc_{n,nn,d,h}$	Net transfer capacities $[MWh_{el}/h]$
$out_{i,n,d,h}$	Power plant outages $[MW_{el}]$
$sc_{i,n,d,h}$	start-up costs $[€/MW_{el}]$
$vc^{FL}_{i,n,d,h}$	Variable generation costs at full load $[€/MWh_{el}]$
$vc^{ML}_{i,n,d,h}$	Variable generation costs at minimum load $[€/MWh_{el}]$
$voll$	Value of lost load $[€/MWh_{el}]$
$wv_{i,n,d,h}$	Water value for hydro reservoirs and long-term storage $[€/MWh_{el}]$
$d^{d+2}_{n,s,h}$	Scenario-specific demand forecast on the two-day-ahead $[MWh_{el}/h]$
$d_{n,d,h}$	Demand forecast on the day-ahead $[MWh_{el}/h]$

Variables

$CL_{i,n,d,h}$	Charging activity for long-term storage $[MWh_{el}/h]$
$CM_{i,n,d,h}$	Charging activity for mid-term storage $[MWh_{el}/h]$
$CURT_{i,n,d,h}$	RES curtailment $[MWh_{el}/h]$
$FLOW_{n,nn,d,h}$	Electricity flow from node n to node nn $[MWh_{el}/h]$
$PCR_{i,n,d,h}$	Primary control reserve $[MW_{el}]$
$SCR^{neg}_{i,n,d,h}$	Negative secondary control reserve $[MW_{el}]$
$SCR^{pos}_{i,n,d,h}$	Positive secondary control reserve $[MW_{el}]$
$SHED_{n,d,h}$	Load shedding $[MWh_{el}/h]$
$SL_{i,n,d,h}$	Storage level of PSP $[MWh_{el}]$
$SU_{i,n,d,h}$	Start-up activity of a generation unit $[MW_{el}]$
TC	Total system costs [€]
$G_{i,n,d,h}$	Electricity generation $[MWh_{el}/h]$
$P^{on}_{i,n,d,h}$	Running capacity $[MW_{el}]$