

Open Energy Transition 组织及其 Solver Benchmark 项目介绍

1. Open Energy Transition 组织 (OET) 介绍

1.1 组织简介

- Open Energy Transition (OET) 是一家成立于 2023 年的非营利性环境咨询、智库和软件开发公司，致力于通过开放数据和开源软件解决方案加速全球向 100% 可再生能源的转型。该组织共同维护着 PyPSA、Linopy 和 atlite 等关键工具，并提供定制化解决方案、软件开发、支持和培训等服务。
- 其工作重点在于通过促进能源规划的透明度、可访问性和协作来产生全球影响。
- OET 的综合模式弥合了研究与实际应用之间的差距，将理论成果转化为可部署工具，并应用于实际能源问题。

1.2 组织结构和主要活动

- OET 拥有由能源系统建模师、软件工程师、分析师和传播专家组成的全球团队，总部位于德国柏林。
- 官方网站为 <https://github.com/open-energy-transition>。
- 在 GitHub 上，OET 维护着 80 个代码库和 3 个项目，拥有 30 名成员。

1.3 Open Source 工具组合和贡献

- OET 通过软件开发、倡导 Open Data 和基础研究，对能源建模领域产生深远影响。
- OET 是多个知名 Open Source 能源规划工具的创建者和/或共同维护者，包括：
 - Python for Power Systems Analysis (PyPSA)。
 - 基于 PyPSA 的欧洲和全球部门耦合能源系统模型 PyPSA-Eur 和 PyPSA-Earth。
 - Linopy 优化接口。
 - 用于将天气数据转换为能源系统数据的 atlite 库。
- OET 的热门代码库包括 Awesome-Electric-Grid-Mapping、Awesome-Grid-Model-Data、solver-benchmark、open-tyndp、MapYourGrid 和 grid-mapping-starter-kit。

1.4 提供的服务和社区参与模式

OET 提供的服务与其非营利使命高度契合，确保 Open Source 工具的长期可用性和可访问性：

- **Customized Solutions:** 提供可复现的报告和模型，赋能客户独立使用和改进工具。
- **Software Development and Maintenance:** 专注于能源系统建模的 Open Source 软件开发和维护，擅长 Python、数据可视化、机器学习和能源领域的优化。
- **Support:** 提供支持合同，确保工具持续维护和改进。
- **Training:** 旨在帮助客户独立并融入社区，培养强大的用户、开发者和维护者网络。

在社区参与方面：

- OET 依赖个人和慈善家的捐赠来维持项目运营和支持社区。
- 致力于在全球范围内，特别是向资源匮乏地区传播能源规划专业知识。
- OET 认为强大的用户、开发者和维护者网络对其项目持续发展至关重要。

2. Solver Benchmark 项目

solver-benchmark 项目是 OET 的核心工作，旨在为能源规划领域的优化求解器提供标准化、开放且可复现的性能比较平台。

2.1 Open Energy Benchmark 项目介绍

- 项目目的：
 - 对能源规划领域的问题进行优化求解器（MILP 和 LP）的基准测试。
 - 系统评估各种优化求解器的性能、计算效率和能力。
 - 提供一个实时、交互式网站 (<https://openenergybenchmark.org/>) 用于分析和可视化基准测试结果。
- 填补空白：
 - 此前缺乏专门针对能源规划中最新且具有代表性问题的基准集。
 - 使求解器开发人员能够针对能源模型的独特特征改进算法。
 - 帮助能源建模人员根据应用、硬件和预算选择最合适的求解器。
- 开源原则：
 - 强调透明度和可复现性，鼓励建模人员提交新的基准实例，并使求解器开发人员能够复现和利用这些基准。
 - 最终目标是加速用于规划和实施全球能源转型以实现可再生能源的关键技术的开发和采用。

2. 技术架构与实施

- GitHub 仓库结构：
 - benchmarks/：包含基准问题（LP/MPS 文件），包括 PyPSA 生成的和来自外部基准的元数据。
 - runner/：包含执行基准测试的核心脚本，如 benchmark_all.sh 和 run_benchmarks.py。
 - results/：存储基准测试运行的输出，包括 metadata.yaml 和 CSV 性能数据文件。
 - website-nextjs/：包含主交互式网站的代码（使用 Next.js）。
 - .github/workflows/：包含 GitHub Actions 工作流文件，用于 CI/CD 自动化。
 - .devcontainer/：开发容器配置文件，确保开发环境一致性。
 - docs/：项目文档。
 - tests/：单元测试和集成测试。
 - .pre-commit-config.yaml 和 ruff.toml：用于代码质量和风格。
- 基准问题：
 - 所有问题均以 LP/MPS 文件形式开放，可从网站下载。
 - 使用 PyPSA 等开源能源模型框架生成，提供 Dockerfile 以实现可复现性。
 - 积极欢迎社区贡献新的基准问题，并为不熟悉 Git/GitHub 的用户提供支持。
- 基准测试执行：
 - runner/benchmark_all.sh 脚本协调运行，提供灵活选项（如运行特定年份、附加结果）。
 - 为每个求解器创建独立的 conda 环境，确保可复现性。
 - runner/run_benchmarks.py 脚本执行求解器，记录运行时间和内存消耗。
 - 支持特定版本的求解器，通常使用每年发布的最新版本（HiGHS、SCIP、CBC、GLPK、Gurobi）。
- 交互式网站：
 - 使用 Next.js 开发，动态处理 CSV 结果和 metadata.yaml 文件。
 - 提供“关键洞察”和“详细结果”部分，用于探索性能数据。
- 可复现性和可扩展性：

- 明确使用 Dockerfiles 和独立的 conda 环境，确保结果的科学有效性和可信度。
- 自动化工作流和简化的贡献过程，使其成为一个高度可扩展和可持续的“活”基准测试平台。

3. 优化求解器性能评估

- **包含求解器：**5 个求解器，包括 4 个开源求解器（GLPK、HiGHS、SCIP、CBC）和 1 个商业求解器（Gurobi）。
 - **CPLEX** (IBM)：用于处理大规模混合整数规划和线性规划问题，支持并行计算。
 - **Gurobi**：广泛用于工业界，专注于提供快速的求解服务，特别擅长大规模的 MILP 问题。
 - **GLPK**：GNU 线性规划工具包，开源，适合处理中小型线性 and 整数规划问题。
 - **MOSEK**：用于优化非线性问题，广泛用于商业和学术领域。
 - **COIN-OR**：开源优化求解器库，支持多种优化技术，包括线性、整数和二次规划。
- **性能指标：**精确测量运行时间、内存消耗，并收集其他指标以确保解决方案质量。
- **数据集：**包含 120 个问题，源自各种能源系统模型，按规模（小型、中型、大型）和“现实”分类。
- **关键发现：**
 - **整体性能：**
 - **小型/中型问题：**GLPK 通常最快，Gurobi 相对较慢。
 - **大型/现实问题：**Gurobi 始终最快，开源求解器慢 20-30 倍。
 - **代表性问题性能：**Gurobi 通常最快，但 4/7 的代表性问题可由至少一个开源求解器（HiGHS 或 SCIP）解决。
- **求解器演变：**
 - Gurobi 在所有年份（2020-2025）保持领先地位。
 - HiGHS 显示出显著改进，版本间加速因子高达 2.5 倍，未解决问题数量保持稳定。
- **开源求解器可行性：**
 - 平台识别了开源求解器可解决的最大 LP 和 MILP 问题，定义了能源建模的“开源可行性边界”。
 - HiGHS 和 SCIP 在解决这些最大问题方面表现突出。
- **影响性能的因素：**
 - **变量数量：**LP/MILP 问题中的变量数量是求解器性能最显著的驱动因素。
 - **模型复杂性：**机组组合（UC）、输电扩展、二氧化碳约束和时间范围等特定特征显著影响求解器效率，有时导致数量级的求解时间增加。
 - 开源求解器在处理 MILP 复杂性方面比商业求解器更吃力。

4. 社区参与和开源生态系统

- **开放贡献模型：**
 - 积极鼓励并促进社区贡献新的基准问题（MPS/LP 文件、YAML 元数据）。
 - 明确为不熟悉 Git/GitHub 的个人提供支持，降低参与门槛。
 - 对包含任何开源求解器持开放态度，并支持社区贡献新的或更新的求解器

版本。

- **GitHub 活动：**
 - **开放问题：**目前有 25 个开放问题，涵盖求解器比较图表清晰度、基准集合空白、网站可用性改进、错误报告和方法论问题（如运行时间测量、对偶间隙容差）。
 - **开放拉取请求：**有 6 个开放拉取请求，表明正在积极开发和集成新功能、错误修复或基准贡献。
 - **主要贡献者：**siddharth-krishna、mattmiltten、danielelerede-oet 和 KristijanFaust-OET 等活跃贡献者。
- **协作研究平台：**
 - 开放问题和方法论讨论表明社区深度投入于基准的准确性、完整性和方法论健全性。
 - 开放贡献模型和活跃的问题跟踪器使其成为一个动态的、协作的研究平台。
 - 通过提供共同的、可复现的测试基础，促进了求解器之间的直接比较和健康竞争。
 - 这种协作环境与 Open Energy Transition (OET) 的使命一致，即“在全球传播能源规划专业知识”。

5. 对能源系统优化和能源转型的影响

- **促进优化算法进步：**
 - 提供“能源规划领域最新且具有代表性问题”的独特测试用例，使求解器开发人员能够精细调整算法。
 - 跟踪求解器性能随时间变化，提供具体、长期反馈（例如，HiGHS 的 2.5 倍加速）。
 - 明确旨在“促进开源优化求解器的发展”，培养开源社区内的竞争和协作。
- **能源建模人员的实用资源：**
 - 提供详细的比较性能数据，帮助建模人员选择“最适合其应用领域、硬件限制和预算的求解器”。
 - 通过清晰展示开源求解器可解决的最大问题，帮助建模人员了解这些工具的计算限制。
 - 开源性质确保建模人员可以复现结果并理解所用方法，促进科学严谨性。
- **促进透明度和开放科学：**
 - AGPL-3.0 许可证和开放 LP/MPS 文件从根本上促进了计算研究的透明度。
 - 严格使用 Docker、Conda 环境和详细执行脚本确保了基准测试运行的完全可复现性。
 - 与能源系统建模中更广泛的开放科学运动完全一致。
 - Open Energy Transition 组织作为牵头方，致力于开发和应用开源工具和方法来推动能源转型。
- **普及高级能源规划能力：**
 - 通过严格的基准测试并突出开源求解器的能力，为学术研究人员、小型组织和财政资源有限的地区普及了高级能源规划能力。
 - 为这些群体提供了清晰的、数据驱动的途径，使其能够在没有高昂财务支出的情况下进行复杂的能源系统建模。

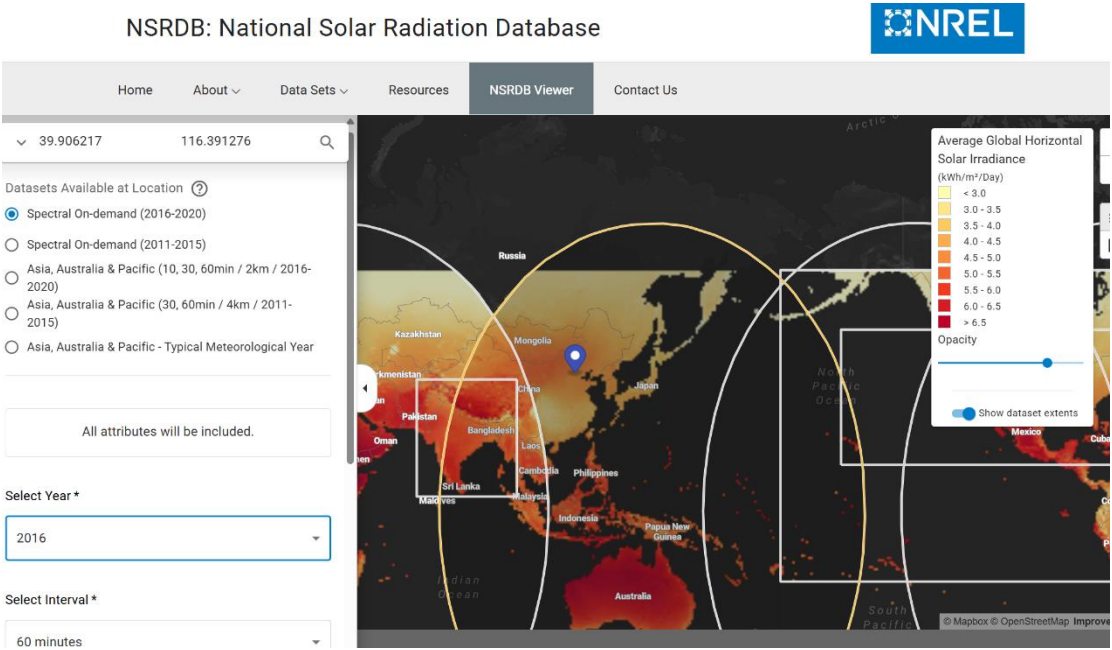
附加任务

1. NREL NSRDB (美国国家可再生能源实验室)

NSRDB 提供两种主要的数据获取方式：网站下载和 API 批量获取。

方法一：通过网站界面下载

- 1. **访问官网：** 前往 [NSRDB 官网](#)。
- 2. **使用数据查看器 (Viewer)：** 登录后，使用网站上的 NSRDB Viewer，通过地图或搜索功能选择感兴趣的地理位置和时间范围。
- 3. **选择数据参数：** 在下载界面，选择所需的数据类型，例如辐照度 (GHI, DNI, DHI)、温度、风速等，并设置时间粒度 (例如 5 分钟、30 分钟或 60 分钟)。
- 4. **下载数据：** 确认后，系统将生成 CSV 数据文件，可以直接下载。



方法二：通过 API 批量获取

对于需要批量或自动化获取数据的研究，可以使用 API。

- 1. **注册获取 API Key：** 在 NREL 开发者网络上注册并申请一个 API Key，这是使用 API 的必要凭证。



[Home](#) [APIs & Documentation](#) [Key Signup](#) [Contact Us](#)

API Key Signup

Use this form to request an API key to use our web services.

Required fields are marked with an asterisk (*).

First Name *

Last Name *

Email *

Website (optional)

How will you use the APIs? (optional)

☐ I have read and agree to the [terms and conditions](#).

[Sign up](#)

2. **查阅 API 文档：** 访问 [NREL 开发者网络](#)，查阅 NSRDB API 的详细文档，了解请求参数（如地理位置坐标、时间范围、数据属性等）和数据格式。
3. **编写脚本：** 根据文档，使用 Python 或其他编程语言编写脚本，通过 API 请求数据。例如，可以使用 requests 库向 API 端点发送 GET 或 POST 请求。

```
import requests

url = "http://developer.nrel.gov/api/climate/ncdb/v2/climate/ncdb-4km-Hourly-CONUS-v1-0-0-RCP8-5-download.json?api_key=yourapikeygoeshere"

payload = "names=2012&leap_day=false&interval=60&utc=false&full_name=Honored%2BUser&email=honored.user%40gmail.com&affiliation=NREL&mailing_list=NREL"

headers = {
    'content-type': "application/x-www-form-urlencoded",
    'cache-control': "no-cache"
}

response = requests.request("POST", url, data=payload, headers=headers)

print(response.text)
```

2. Kaggle

Kaggle 是一个数据科学社区，可以通过其平台直接下载用户共享的数据集。

1. **访问 Kaggle 网站：** 前往 [Kaggle 官网](#)。
2. **搜索数据集：** 在搜索栏中输入关键词，例如“PV power”等。

3. **选择数据集：**从搜索结果中浏览并选择一个感兴趣的数据集(Datasets)。例如，“Solar Power Generation Data”。

← pv power

<> Notebooks 331 Datasets 73 Topics 13 Comments 6 Competitions 6 People 2 Hugging Face Models 1

Filter by 73 Results Relevance ▾

DATE

☐ Last 90 days 7

VIEWED BY YOU

☐ Viewed 1

☐ Not Viewed 72

CREATOR

☐ You 0

☐ Others 73

- Solar Power Generation Data**
Dataset - 1y ago · by Afroz
For example, according to the accurate forecast, PV system operators can balance the consumption of power. 4,176 downloads
- Hourly Power Generation data from Solar PV Plant**
Dataset - 1y ago · by Afroz
From 2017 September to 2024 February 708 downloads
- Global Energy Consumption & Renewable Generation**
Dataset - 4y ago · by James Arthur
main renewable energy sectors : Hydro, Wind, Biofuel, Solar PV, and Geothermal. 8,341 downloads
- Solar energy power generation dataset**
Dataset - 4y ago · by Meshua
Accurate forecasting of the power output of PV systems in a short term is of great importance for daily 5,131 downloads
- Electricity consumption UK 2009-2024**
Dataset - 7mo ago · by Alberto Vidal
-ve signifies export power out from GB; +ve signifies import power into GB. Measured in MW. 5,473 downloads
- California Renewable Production 2010-2018**
Dataset - 8y ago · by Sid Ghodke
Hourly Power Production in California 1,824 downloads

4. **下载文件：**进入数据集页面后，通常会看到一个“Download”按钮。点击该按钮即可下载包含 CSV 或其他格式数据文件的压缩包。

≡ kaggle

+ Create

Home Competitions Datasets Models Benchmarks Code Discussions Learn More

Your Work

VIEWED

Global Energy Consu... Gemma Solar Powe Generatio... View Active Events

Search

Solar Powe Generation Data 81 Code

Data Card Code (12) Discussion (1) Suggestions (0)

Solar Power Plant Data.csv (408.02 kB) Download 8 of 8 columns ▾

Detail Compact Column

About this file [Suggest Edits](#)

This file does not have a description yet.

Δ Date-Hour(NMT)	# WindSpeed	# Sunshine	# AirPressure	# Radiation
8760 unique values				
01.01.2017-00:00	0.6	0	1003.8	-7.4
01.01.2017-01:00	1.7	0	1003.5	-7.4
01.01.2017-02:00	0.6	0	1003.4	-6.7
01.01.2017-03:00	2.4	0	1003.3	-7.2
01.01.2017-04:00	4	0	1003.1	-6.3
01.01.2017-05:00	1.4	0	1003.1	-6.8
01.01.2017-06:00	1.4	0	1003.7	-7
01.01.2017-07:00	1.3	0	1003.9	-7

3. GEFCom (全球能源预测竞赛)

GEFCom 竞赛已结束，但其历史赛题数据通常会公开发布。获取这些数据需要找到竞赛组织者发布的官方链接。

1. **搜索历史竞赛资料：**使用搜索引擎搜索“GEFCom2014”、“GEFCom data download”等关键词。

2. **定位官方发布链接：**搜索结果通常会指向竞赛组织者或相关学术论文的博客、GitHub 页面或官方存档。例如，GEFCom2014 的数据曾在组织者 Tao Hong 博士的博客上提供下载链接。



3. **下载数据：** 找到数据下载链接后，通常是一个 Dropbox 或其他网盘的分享链接，点击即可下载包含历史赛题数据的压缩包。

GEFCom2014.zip / **GEFCom2014 Data**

文件 编辑 查看 帮助

提取所有

下载

名称	大小
GEFCom2014-E_V2.zip	2.24 MB
GEFCom2014-L_V2.zip	2.48 MB
GEFCom2014-P_V2.zip	3.18 MB
GEFCom2014-S_V2.zip	35.03 MB
GEFCom2014-W_V2.zip	77.39 MB
Provisional_Leaderboard_V2.xlsx	191.34 KB
READ ME_V2.txt	389 B