

Digital Discovery

Accepted Manuscript

This article can be cited before page numbers have been issued, to do this please use: P. Seyitdanlioglu, *Digital Discovery*, 2026, DOI: 10.1039/D6DD00247A.



This is an Accepted Manuscript, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this Accepted Manuscript with the edited and formatted Advance Article as soon as it is available.

You can find more information about Accepted Manuscripts in the [Information for Authors](#).

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard [Terms & Conditions](#) and the [Ethical guidelines](#) still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this Accepted Manuscript or any consequences arising from the use of any information it contains.

ARTICLE

A reproducible Python workflow for absorber–light-source spectral matching: overlap-calculator

Pinar Seyitdanlioglu^{*a}Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

The spectral compatibility between an organic absorber and the illumination source is an important but often under-quantified descriptor in computational screening of organic photovoltaic materials, particularly for indoor applications where light sources have narrow and source-dependent emission profiles. Here, we introduce overlap-calculator, an open-source Python workflow for reproducible batch analysis of spectral overlap between molecular absorption spectra and solar or indoor reference light sources. The workflow accepts Gaussian TD-DFT output files and tabular UV-vis spectra in CSV or Excel formats within a common manifest-driven pipeline. TD-DFT transitions are reconstructed into continuous absorption profiles using Gaussian and Lorentzian broadening, converted into absorbance through a Beer–Lambert treatment, and compared with AM1.5G, CIE LED, fluorescent, or user-supplied spectra. The resulting descriptors include absorbed flux, absorbed fraction, and max-normalised shape overlap. The workflow is demonstrated using five previously generated TD-DFT OPV candidate absorbers and eight public organic UV-vis spectra used as spreadsheet-input examples. The TD-DFT case study illustrates automated transition parsing, spectral reconstruction, and source-dependent ranking, whereas the spreadsheet-input case demonstrates that tabular UV-vis data can be processed through the same descriptor-generation pipeline. By converting a previously script-based interpretation step into a documented and reusable workflow with command-line, Python-library, HTTP-API, Docker, structured-output, and per-run provenance-manifest support, overlap-calculator provides a practical tool for transparent spectral-compatibility analysis in data-driven optoelectronic materials research.

Introduction

Organic photovoltaics (OPVs) are attractive for emerging energy-harvesting applications because molecular structure can be tuned to control optical absorption, frontier orbital energies, excited-state character, and processing behaviour.^{1–3} In addition to outdoor solar conversion, OPVs are increasingly discussed for indoor and low-irradiance environments, where lightweight, flexible, and solution-processable devices could power distributed sensors or low-power Internet-of-Things devices.^{4,5} This application space changes the screening problem. Indoor and low-light photovoltaic operation differs from outdoor operation in ways that motivate, rather than strictly require, spectrally tailored absorbers. The photon flux available from common artificial sources (warm-white LEDs, fluorescent lamps) is concentrated in the visible window and is, per unit area, two to three orders of magnitude smaller than AM1.5G, which shifts the radiative-limit optimum towards wider-gap absorbers and amplifies the penalty of any mismatch between the absorption profile and the source spectrum. Recent indoor organic and perovskite cells have exceeded ~28–30% power-conversion efficiency under indoor illuminants precisely by exploiting this spectral match, so a quantitative descriptor of

source–absorber overlap is a useful pre-synthesis screen; however, because a sufficiently broad and inexpensive absorber may also be a viable route, the criterion is best framed as a design advantage rather than a strict requirement.^{6–14} Under indoor lighting, the relevant illumination source is no longer only a broad solar spectrum but also a source-specific emission profile, such as a phosphor-converted white LED or a fluorescent lamp. Therefore, an absorber that appears promising under AM 1.5G illumination may not be equally suitable under a narrow or structured indoor spectrum.^{13–15}

Computational materials screening commonly evaluates descriptors such as the HOMO–LUMO gap, vertical excitation energy, oscillator strength, open-circuit voltage estimates, light-harvesting efficiency, reorganization energy, and charge-transfer metrics.^{16–18} In TD-DFT-based screening studies, the spectral suitability of candidate photovoltaic materials is commonly assessed using absorption maxima, oscillator strengths, light-harvesting efficiencies, and simulated UV–Vis absorption profiles.^{19–21} These descriptors remain essential, yet they do not by themselves quantify how much of a given light-source spectrum overlaps with the wavelength-dependent absorbance of a candidate material.^{22,23} The absorption maximum alone is particularly insufficient for indoor OPVs: two molecules with similar peak positions may show different spectral widths, line shapes, and degrees of overlap with LED or fluorescent emission bands. Conversely, molecules with different absorption maxima may exhibit comparable spectral

^a Department of Chemistry, Faculty of Science, Hacettepe University, Ankara, Türkiye.
E-mail: pinarseyit@gmail.com



compatibility when their broadened absorption profiles cover the same high-intensity region of the light source.^{22, 24, 25}

Although useful for early-stage screening, such descriptor-level or visual interpretation can limit reproducibility and make it difficult to compare TD-DFT-derived transition data and spreadsheet-based spectra on a common quantitative basis.^{26, 27} In practice, spectral compatibility is often evaluated by inspecting overlaid absorption and light-source spectra or by applying small workflow-specific calculations. While useful, such approaches can obscure processing choices such as spectral interpolation, normalization, broadening, and integration limits.^{22, 26, 28} This is a natural starting point during exploratory research, but it creates reproducibility challenges. First, theoretical TD-DFT spectra and experimental UV-Vis spectra are often handled with different scripts or assumptions. Second, the choice of broadening function, wavelength grid, concentration, path length, normalization, and light-source data can remain implicit. Third, the resulting plots and tables are rarely generated from a single documented workflow that can be rerun on a new molecule set. These issues are especially important when spectral compatibility is used to rank candidates or justify molecular design decisions.

To address this need, the present work introduces overlap-calculator as a reproducible Python workflow for batch spectral-overlap analysis between molecular absorption spectra and solar or indoor light sources. The implementation supports multiple solar and indoor light sources, Gaussian TD-DFT outputs, spreadsheet UV-vis spectra, structured exports, command-line and Python API access, Docker deployment, provenance-aware output generation, and publication-quality visualization. Rather than replacing established electronic-structure or workflow-management software, overlap-calculator provides a focused post-processing layer that standardizes the transformation of absorption information into light-source-dependent overlap descriptors. This makes the analysis easier to audit, repeat, and extend.

Workflow definition of overlap-calculator

The design of overlap-calculator follows a workflow-centred strategy in which the scientific operation of spectral-overlap analysis is separated from the technical details of file discovery, parsing, interpolation, spectral broadening, absorbance conversion, light-source handling, and output generation. The workflow was developed to support a recurring post-processing task in photovoltaic materials screening: comparing molecular absorption spectra with solar or indoor illumination spectra on a common quantitative basis (Fig. 1).

The package accepts both theoretical and experimental absorption data within the same analysis run. Theoretical inputs are Gaussian TD-DFT output files,²⁹ from which excited-state wavelengths and oscillator strengths are extracted. Experimental inputs are tabular UV-Vis spectra supplied as files or Excel workbooks. By treating both input families as first-class

data sources, the workflow allows computed and measured spectra to be processed using the same downstream descriptors. This is particularly useful when TD-DFT predictions are compared with available UV-Vis measurements or when a computational screening study is extended toward experimental validation.

The workflow is organised into six main stages. First, input files are resolved through a JSON manifest that can be generated automatically from a directory of raw files or edited manually by the user. Second, theoretical and experimental inputs are parsed using input-specific loaders. Gaussian TD-DFT files are processed to extract discrete excited-state information, whereas tabular spectra are read as wavelength-dependent absorbance data. Third, TD-DFT transitions are reconstructed as continuous absorption spectra on a shared wavelength grid using Gaussian and Lorentzian line-shape functions. Fourth, the absorption information is converted into absorbance through the Beer-Lambert relationship. Fifth, the resulting absorbance spectra are integrated against selected reference light-source spectra. Finally, the workflow exports structured result tables, descriptor summaries, skip reports, timing information, and publication-quality plots (Fig. 1). The main components of the workflow and their scientific roles are summarized in Table 1. The exact mixed input manifest used in the demonstration is provided in Table S2 of the ESI.

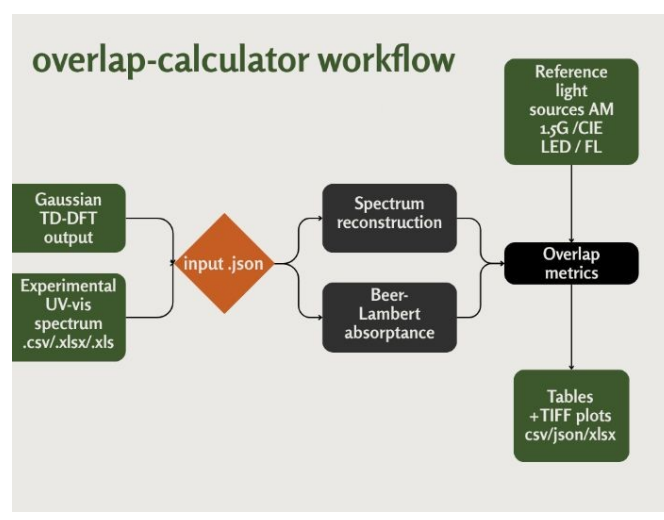


Fig. 1 Conceptual workflow of overlap-calculator.

Input families and manifest-driven analysis

overlap-calculator treats theoretical and experimental absorption data as two first-class input families. The theoretical path accepts Gaussian TD-DFT output files with .out or .log extensions. The experimental path accepts comma-separated values files and Excel workbooks, including multi-sheet and multi-series datasets. This design allows computed and measured spectra to be processed within a single analysis run, which is useful when computational screening results are compared with available UV-Vis measurements or when



theoretical predictions are connected to experimental validation.

This manifest-driven design avoids hard-coded file paths in the analysis code and makes the analysed dataset explicit. It also improves traceability because each result row can be linked back to a specific raw file, sample identifier, sheet, and spectral series. As a result, the same analysis can be rerun on the same dataset or extended to a new molecule set without modifying the workflow code.

Table 1 Main components of the overlap-calculator workflow and their scientific roles in reproducible spectral-overlap analysis.

Component	Role in the workflow	Scientific purpose
input.json manifest	Defines sample identifiers, file paths, input type, series name, and sheet name.	Makes the analysed dataset explicit and rerunnable.
TD-DFT parser	Extracts excited-state wavelengths, oscillator strengths, and route metadata from Gaussian outputs.	Transforms discrete excited-state information into spectral inputs.
Experimental loader	Reads wavelength and absorbance columns from CSV or Excel files.	Allows measured UV-vis spectra to be processed with the same overlap descriptors.
Light-source loader	Loads AM1.5G, CIE LED/FL, or custom two-column spectra.	Enables source-specific spectral compatibility analysis.
Descriptor exporter	Writes results, summary tables, skip reports, timings, and plots.	Supports ranking, troubleshooting, and publication-ready visualisation.
Calibration block (optional)	Applies $E_i^{cal} = aE_i^{calc} + b, f_{cal} = a * f$	Per-band overrides before broadening; identity bit-identical.

A mixed theoretical-experimental manifest can contain Gaussian TD-DFT outputs and spreadsheet-based UV-Vis spectra in the same run:

```
[
  {
    "input_type": "theoretical",
    "sample_id": "B1_td",
    "source_path": "files/B1_td.out",
    "series_name": null,
    "sheet_name": null
  },
  {
    "input_type": "experimental",
    "sample_id": "dye_panel_channel_A",
    "source_path": "files/dye_panel.xlsx",
    "series_name": "channel_A",
    "sheet_name": "raw"
  }
]
```

Although overlap-calculator is presented primarily through its command-line interface, the underlying analysis is exposed as an importable Python library: the spectrum-loading, reconstruction, overlap-integral, and descriptor routines are public, typed functions that can be called directly from a notebook or a larger pipeline. This positions the tool for orchestration from established workflow-management systems for computational materials science—AiiDA, jobflow, and pyiron — in which it would act as a calculator node that consumes upstream TD-DFT results and absorption spectra.³⁰⁻³² Native plug-ins for these engines are not implemented in the present release and are noted as integration targets; the structured input contract and per-row provenance are also deliberately compatible with the recently proposed Python Workflow Definition exchange format, which we identify as future work for cross-engine portability.³³

From oscillator strengths to broadened extinction spectra

For theoretical inputs, the workflow converts discrete TD-DFT transitions into continuous absorption spectra on a shared wavelength grid. Each excited state is represented according to its oscillator strength and broadened using both Gaussian and Lorentzian line-shape functions. Reporting both broadening models is useful because spectral width, tailing behaviour, and the resulting overlap with a light-source spectrum can depend on the selected reconstruction model, particularly when the illumination profile contains narrow LED or fluorescent emission bands.

For each TD-DFT transition i , the excitation wavelength is converted to a wavenumber scale and represented by a normalised line profile. The wavelength grid is converted according to

$$\tilde{\nu} = 10^7 / \lambda \quad (\lambda \text{ in nm; } \tilde{\nu} \text{ in cm}^{-1}) \quad (1)$$

The Gaussian and Lorentzian profiles are then defined as:

$$g_i(\tilde{\nu}) = [1 / (\sigma\sqrt{2\pi})] \exp\{- (\tilde{\nu} - \tilde{\nu}_i)^2 / (2\sigma^2)\} \quad (2)$$

$$l_i(\tilde{\nu}) = \left(\frac{1}{\pi}\right) \left[\frac{\gamma}{(\tilde{\nu} - \tilde{\nu}_i)^2 + \gamma^2}\right] \quad (3)$$

$$\gamma = \sqrt{2\ln 2}\sigma \approx 1.1774\sigma \quad (4)$$

Here, σ denotes the Gaussian standard deviation. For the Lorentzian profile, the half-width-at-half-maximum γ is defined as $\sqrt{2\ln 2}\sigma$, ensuring that the Gaussian and Lorentzian line shapes have the same FWHM for a given --sigma-ev value. This choice removes an implicit width mismatch between the two broadening models, so that the reported `delta_*` columns primarily reflect differences in tail behaviour rather than differences in spectral width. The molar extinction coefficient is reconstructed as a weighted sum over oscillator strengths:

$$\varepsilon(\lambda) = 2.315 \times 10^8 \sum_i f_i p_i(\tilde{\nu}(\lambda)) \quad [\text{M}^{-1} \text{cm}^{-1}] \quad (5)$$



The prefactor $P = 2.315 \times 10^8 \text{ M}^{-1}\text{cm}^{-2}$ in equations 2 and 5 is the inverse of the integrated-intensity relation $f_i = 4.319 \times 10^{-9} \int \epsilon_i(\tilde{\nu}) d\tilde{\nu}$ (with $\tilde{\nu}$ in cm^{-1} and ϵ_i in $\text{M}^{-1}\text{cm}^{-1}$), distributed over a line shape normalised to unit area on the wavenumber axis. Taking P outside the integral is exact in the narrow-line limit $\sigma \ll \tilde{\nu}_i$ and otherwise incurs a fractional error of order $(\sigma/\tilde{\nu}_i)^2$. For the default broadening $\sigma = 0.30 \text{ eV}$ and the 200–800 nm window targeted here the raw $(\sigma/\tilde{\nu})^2$ scale reaches $\sim 3.8\%$ at the red edge ($\lambda = 800 \text{ nm}$, $\tilde{\nu} = 12\,500 \text{ cm}^{-1}$), but because the leading $\tilde{\nu}$ -resolved correction integrates out by symmetry for a symmetric line shape, the residual error on the band area / peak stays below $\sim 2\%$ over 200–800 nm well within the uncertainty already introduced by the phenomenological σ . The constant-prefactor form is therefore retained as the default for visible–ultraviolet screening, and a frequency-resolved option is provided for users extending the workflow into the near-infrared, where a $\tilde{\nu}$ -resolved $P(\tilde{\nu})$ must be evaluated inside the integral.^{34, 35} The wavelength grid, broadening width, and numerical integration window are user-controllable. The frequency-resolved option is intended for near-infrared or lower-energy applications.

The core reconstruction step can be represented as follows:

```
import numpy as np

def gaussian_profile(nu, nu_i, sigma):
    return np.exp(-0.5 * ((nu - nu_i) /
sigma) ** 2) / (
sigma * np.sqrt(2 * np.pi)
)

def lorentzian_profile(nu, nu_i, sigma):
    return (1 / np.pi) * sigma / ((nu -
nu_i) ** 2 + sigma**2)

def reconstruct_extinction(wavelength_nm,
transitions, sigma_cm, profile):
    nu = 1e7 / wavelength_nm
    epsilon = np.zeros_like(wavelength_nm,
dtype=float)

    for transition in transitions:
        nu_i = 1e7 /
transition["wavelength_nm"]
        f_i =
transition["oscillator_strength"]
        epsilon += f_i * profile(nu, nu_i,
sigma_cm)

    return 2.315e8 * epsilon
```

The use of two-line shapes is not intended to imply that either function universally reproduces an experimental UV-Vis spectrum. Instead, it provides a transparent sensitivity check. If the ranking of candidate absorbers or light sources remains

stable across Gaussian and Lorentzian reconstructions, the conclusion is less likely to be an artefact of the selected spectral reconstruction model. The user-facing parameter is given in eV; the equation uses σ in cm^{-1} via $\sigma_{\text{cm}^{-1}} = 8065.544\sigma_{\text{eV}}$. Additional implementation details for spectral reconstruction and absorbance conversion are provided in Section S5 of the ESI.

Beer-Lambert absorbance conversion

For experimental inputs, the user-provided spectral signal is treated as absorbance, $A(\lambda)$, and is processed on the same wavelength grid used for theoretical spectra. CSV files and Excel workbooks are parsed to identify a wavelength column and one or more numeric signal columns. Each numeric signal column is treated as an independent spectrum, allowing multi-series UV-Vis datasets to be analysed without preparing separate input files for each sample.

Both theoretical and experimental spectra are converted into wavelength-dependent absorbance before overlap descriptors are calculated. For TD-DFT-derived spectra, absorbance is obtained from the reconstructed molar extinction coefficient using the Beer-Lambert relationship:

$$A(\lambda) = \epsilon(\lambda)cL \quad (6)$$

$$\alpha(\lambda) = 1 - 10^{-A(\lambda)} \quad (7)$$

where c is the reference concentration and L is the optical path length. For experimental spectra, the measured absorbance is inserted directly into $\alpha(\lambda) = 1 - 10^{-A(\lambda)}$. This common treatment places TD-DFT-derived and spreadsheet-based spectra on the same numerical basis before they are compared with light-source spectra. The reference concentration and path length are configurable parameters rather than fixed assumptions, and the selected values are written to the output tables so that the assumptions used in the analysis remain visible and reproducible.

Experimental spectra are treated as absorbance directly and passed through the same absorbance definition:

```
def extinction_to_absorbance(epsilon,
concentration_molar=1e-5, path_cm=1.0):
    absorbance = epsilon *
concentration_molar * path_cm
    return 1 - 10 ** (-absorbance)
```

```
def
measured_absorbance_to_absorbance(absorba
nce):
    return 1 - 10 ** (-absorbance)
```

The runtime Beer-Lambert reference state and wavelength-grid settings used in the manuscript are listed in Section S3 of the ESI.



Light-source-dependent overlap descriptors

The workflow compares each molecular absorptance spectrum with one or more reference light-source spectra. Built-in light sources include the AM 1.5G solar spectrum and representative CIE LED and fluorescent illuminants.³⁶⁻³⁹ Users can also supply custom light-source spectra as two-column files containing wavelength and intensity values. This flexibility allows the same molecular dataset to be evaluated under outdoor solar irradiation, standard indoor illuminants, or public experimental lamp spectra.

For each sample and light source, the workflow computes the following integrals over the selected wavelength grid:

$$absorbed_{flux} = \int \alpha(\lambda)I(\lambda)d\lambda \quad (8)$$

$$light_{flux_total} = \int I(\lambda) d\lambda \quad (9)$$

$$absorbed_{fraction} = \frac{absorbed_{flux}}{light_{flux_total}} \quad (10)$$

$$shape_{overlap} = \frac{[\int \min(\hat{\alpha}(\lambda), \hat{I}(\lambda)) d\lambda]}{[\int \hat{I}(\lambda) d\lambda]} \quad (11)$$

Here $I(\lambda)$ is the light-source spectral distribution and hats indicate max-normalised profiles. `absorbed_fraction` reports the fraction of the available spectral source captured by the absorptance profile within the grid, while `shape_overlap` focuses on the relative matching of spectral shapes. The latter is particularly useful for CIE LED and fluorescent spectra, which are provided as relative spectral power distributions.

Once the molecular absorptance and the light-source spectrum are defined on the same wavelength grid, the absorbed flux, absorbed fraction, and shape-overlap descriptors are calculated numerically:

```
import numpy as np

def overlap_descriptors(wavelength_nm,
                        absorptance, light_intensity):
    total_flux = np.trapz(light_intensity,
                          wavelength_nm)
    absorbed_flux = np.trapz(absorptance *
                              light_intensity, wavelength_nm)

    absorbed_fraction = absorbed_flux /
total_flux if total_flux > 0 else np.nan

    absorptance_norm = absorptance /
np.nanmax(absorptance)
    light_norm = light_intensity /
np.nanmax(light_intensity)

    shape_overlap = np.trapz(
        np.minimum(absorptance_norm,
                    light_norm),
        wavelength_nm
    ) / np.trapz(light_norm, wavelength_nm)
```

```
return {
    "light_flux_total": total_flux,
    "absorbed_flux": absorbed_flux,
    "absorbed_fraction":
absorbed_fraction,
    "shape_overlap": shape_overlap,
}
```

View Article Online

The light-source definitions and descriptor equations used in the demonstration are summarised in Sections S6 and S7 of the ESI.

Command-line interface

The command-line interface is designed around two main steps. First, `generate-input` scans a directory and writes a manifest. Second, `analyze` processes the manifest and writes tables and plots to an output directory. This separation is useful for publication workflows because the manifest can be archived together with the raw inputs and output tables. The command-line execution is shown below:

```
overlap-calculator generate-input --files-dir
input/files --out input/input.json
```

```
overlap-calculator analyze --input
input/input.json --out output \
--sigma-ev 0.30 --wl-min 200 --wl-max 800
--num-points 10000 \
--concentration-m 1e-5 --path-cm 1.0 \
--default-light-sources
AM15G, LEDB2, LEDB3, LEDB4, CIEFL10
```

The most relevant runtime parameters are the broadening width, wavelength range, number of grid points, reference concentration, path length, default light sources, custom light-source files, plot generation flag, and logging configuration. This two-step structure allows the dataset description to be inspected and archived before the numerical analysis is executed, and it makes it possible to rerun the same analysis with different settings without changing the raw input files.

HTTP API

For deployment or integration with a web interface, the package also exposes a lightweight HTTP API. The `/health` endpoint provides a liveness check, while `/analyze` accepts multipart uploads containing mixed theoretical and experimental files, optional custom light sources, and runtime parameters. The response is a ZIP archive containing the same file structure generated by the command-line workflow. This makes the tool usable both as a local research script and as a reproducible service:



```
curl -X POST http://localhost:8000/analyze
\
-F "files=@input/files/theoretical.out"
\
-F
"files=@input/files/experimental.xlsx" \
-F "sigma_ev=0.25" \
-F "default_light_sources=AM15G,LEDB4" \
--output analysis_outputs.zip
```

Choice of broadening

The single user-facing width parameter σ , used to convert TD-DFT stick spectra into a continuum, is an explicitly phenomenological screening parameter. For the Lorentzian reconstruction, the half-width-at-half-maximum is defined as $\gamma = \sqrt{(2 \ln 2)\sigma}$, ensuring that the Gaussian and Lorentzian line shapes have matched FWHM values. It is identical for every transition, is fully under user control (default $\sigma = 0.30$ eV), and is not derived from any underlying physical width: no reorganisation energy is partitioned into the band shape, no Huang–Rhys factors are evaluated, and no Duschinsky mixing is included. This convention is retained because the workflow is a descriptor calculator built on top of an absorption profile, not a band-shape simulator, and because a single transparent parameter is more useful for screening than a hidden composite of unverified assumptions; to let users audit robustness, the pipeline emits Gaussian, Lorentzian, and `delta_*` columns side by side and supports a σ sweep. When quantitative band shapes are required — vibronic progressions, mirror-image symmetry with fluorescence, accurate 0–0 placement, or low-temperature spectra — this convenience path is not the recommended route. Two physically grounded alternatives feed the same descriptor pipeline. First, a Marcus-type Gaussian width, $\sigma_i = \sqrt{(2\lambda_i k_B T)}$, is a built-in broadening mode (`--sigma-mode marcus-hush` with `--reorganization-ev` and `--temperature-k`; library entry point `marcus_hush_sigma_ev`) that assigns the classical high-temperature width per transition before broadening; at $\lambda = 0.30$ eV and $T = 298.15$ K this gives $\sigma = 0.124$ eV (FWHM 0.292 eV), the classical limit, so for organic vibrational baths with $\hbar\omega \gg k_B T$ it is a first-order estimate. Second — recommended for production work — a fully vibronic Franck–Condon or Franck–Condon plus Herzberg–Teller spectrum computed with Gaussian (Freq=FC), DUSHIN, or FCclasses, exported on a wavelength grid and ingested through the tabular-input branch.^{40–44}

Output structure

The output directory is divided into tables and plots. The tables directory contains results in CSV, JSON, and XLSX formats, a descriptor summary, timing information, and a skipped-input report if applicable. The plots directory contains per-sample absorbance spectra, light-source spectra, overlap plots, and multi-sample overlays. All TIFF plots are intended for

manuscript preparation and are produced in both absolute and max-normalised forms. DOI: 10.1039/D6DD00247A

The main output table contains one row per sample and light-source pair, with separate columns for Gaussian and Lorentzian broadening. The most important descriptors are the peak absorbance wavelength, peak absorbance, absorbed flux, absorbed fraction, and shape overlap (Table 2). Additional delta columns report the difference between Gaussian and Lorentzian metrics. Because the tables and plots are generated from the same documented workflow, the numerical descriptors and visual outputs remain consistent with one another.

Table 2 Main descriptors exported by overlap-calculator and their suggested use.

Descriptor	Definition or interpretation	Suggested use
absorbed_flux	Integral of absorbance multiplied by the light-source spectrum.	Quantifies source-weighted absorbed light; unit depends on the source spectrum.
absorbed_fraction	absorbed_flux divided by total light-source flux over the grid.	Compares how much of the selected source is absorbed by a sample.
shape_overlap	Integral comparator using max-normalised absorbance and light-source shapes.	Useful when light-source data are relative, especially for LED and fluorescent spectra.
delta metrics	Gaussian minus Lorentzian descriptor values.	Checks sensitivity to the broadening model.
descriptor_summary	Condensed per-sample table across light sources.	Starting point for ranking candidate absorbers.

Two grouped ranking tables, `ranking_by_light_source_<metric>` (samples ordered within each light source) and `ranking_by_sample_<metric>` (light sources ordered within each sample), are emitted in CSV/JSON/XLSX for all four-overlap metrics (gaussian/lorentzian × absorbed_fraction/shape_overlap). Companion bar charts accompany each table under `plots/ranking/`. The generated output files and ranking-table families are listed in Section S8 of the ESI.

Reproducibility and deployment

The workflow can be installed through a standard Python environment, Conda environment, or Docker image. Docker deployment is useful when the analysis is shared with collaborators who do not have the same local Python environment. For archival purposes, the final publication should provide a versioned GitHub release and a Zenodo DOI, together with the input manifest, example data, and output tables used in the manuscript. Command-line, API, Docker, software-environment, and test commands are provided in Sections S10–S13 of the ESI.

The design of overlap-calculator is aligned with the FAIR Principles for Research Software (FAIR4RS), which adapt the FAIR Guiding Principles from data to executable software.^{45,46}



The source code is openly developed under a permissive licence, every tagged release is archived with a persistent Zenodo DOI, and the dependency surface is reproducible through a pinned environment and a versioned container image, supporting the findable, accessible, and reusable dimensions; interoperability is addressed by reading widely supported tabular inputs and emitting structured, schema-stable outputs. Throughout this manuscript we use the term traceability in the sense of computational provenance: every analyze run writes a `run_manifest.json` at the output root recording the software version, git commit, UTC timestamp, the full resolved parameter set, the light-source set, and one entry per (source, sample, series, sheet) with the SHA-256 of each input file; each descriptor row additionally carries `prefactor_mode`, `sigma_mode`, and `software_version`, and is linked via the input manifest to its originating file, sheet, series, sample, and spectrum so that a third party can re-execute the analysis and reproduce the descriptors from the archived artefacts alone.^{27,47}

Demonstration workflows

Rationale of the demonstrations

The manuscript uses two complementary demonstrations to separate the scientific application from the input-format validation. The demonstration set was designed to evaluate both input branches of overlap-calculator. Five Gaussian TD-DFT output files, generated previously by the author and reused here only as representative calculated-spectrum inputs,⁴⁸ were included to demonstrate automated transition parsing, spectral reconstruction, Beer–Lambert absorptance conversion, and light-source-dependent descriptor generation. These files were not introduced as a new TD-DFT dataset, and the present analysis does not re-evaluate the molecular systems from the original study. Instead, the examples serve solely to show how overlap-calculator processes Gaussian TD-DFT outputs. The resulting broadened spectra were then compared with AM1.5G, LED, and fluorescent illumination spectra (Fig. 2–4). In parallel, eight experimentally reported organic UV-vis spectra were arranged as spreadsheet inputs to evaluate the tabular-data branch of the workflow. These spectra were not intended as an experimental validation of the TD-DFT candidates; rather, they demonstrate that non-Gaussian UV-vis spectra can be ingested, converted to absorptance, and analysed using the same light-source-dependent descriptor pipeline. It should be emphasised that the PhotochemCAD-derived molecules used in the spreadsheet-input demonstration are not proposed here as OPV candidate absorbers. They were selected only as publicly available organic UV-vis spectra with diverse absorption profiles, in order to test whether overlap-calculator can ingest tabular experimental spectra, convert them into absorptance profiles, and process them through the same descriptor-generation pipeline used for TD-DFT-derived inputs. Therefore, this part of the demonstration validates the

experimental/spreadsheet input branch of the workflow rather than the photovoltaic relevance of these particular molecules. The demonstration inputs are not a representative chemical space: the five TD-DFT files are a small B3LYP/6-31G donor–acceptor set, and the PhotochemCAD-derived spreadsheet spectra were chosen as well-characterised, strongly absorbing visible chromophores. They are used solely to exercise the two input branches of the workflow and are not proposed as a screening corpus or as OPV candidate absorbers.

Together, the demonstrations address three practical questions: (i) how Gaussian and Lorentzian broadening affect TD-DFT-derived overlap descriptors, (ii) whether candidate rankings change between outdoor and indoor light sources, and (iii) whether non-Gaussian UV-vis spectra arranged in Excel format can be converted into the same descriptor tables and plots as Gaussian output files. Details of the theoretical and spreadsheet-input branches used in the demonstration are provided in Sections S2–S4 of the ESI.

For each candidate absorber, Gaussian TD-DFT output files are parsed to extract excited-state wavelengths and oscillator strengths. The spectra are reconstructed over the same wavelength grid using both broadening models. The resulting absorptance profiles are then integrated against AM 1.5G, LEDB2, LEDB3, LEDB4, and CIEFL10 spectra. This produces a matrix of light-source-dependent descriptors suitable for ranking and visualisation.

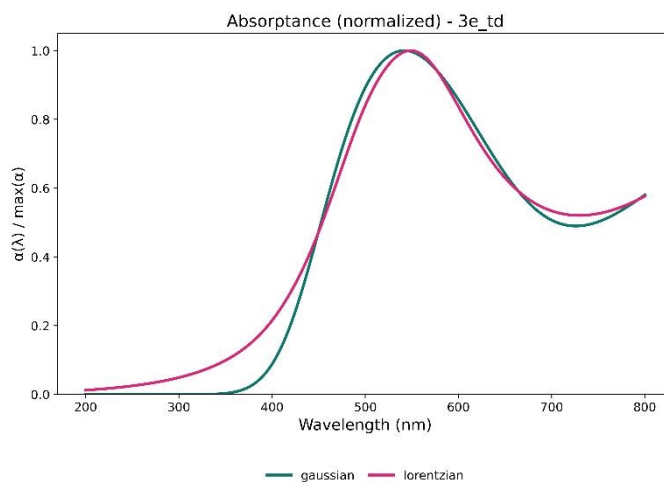


Fig. 2 Representative TD-DFT-derived absorption profile reconstructed using Gaussian and Lorentzian broadening.

TD-DFT-derived OPV screening case

We emphasise that the workflow is method-agnostic: the parser accepts Gaussian TD-DFT output files regardless of the functional and basis set used, and the tabular branch accepts



any wavelength-gridded spectrum, whether computed at a higher level of theory, simulated vibronically, calibrated empirically, or measured experimentally. The B3LYP/6-31G level of the five demonstration files reflects the legacy inputs reused here only to demonstrate parsing, spectral reconstruction, broadening, and descriptor generation; it is not a recommendation for new excited-state calculations and does not restrict the applicability of the workflow. For new electronic-structure inputs, current excited-state TD-DFT benchmark studies for organic chromophores generally show that the accuracy depends strongly on the exchange–correlation functional, the basis set, and the charge-transfer character of the transition; range-separated hybrids such as CAM-B3LYP or ω B97X-D, as well as modern global hybrids such as M06-2X and PBE0, are therefore often preferred over uncalibrated conventional hybrids.^{49–52} To address the systematic errors documented in those benchmarks, overlap-calculator (v1.1.0 onward) ships a calibration block, supplied via `--calibration PATH (CLI)` or `Calibration / load_calibration / apply_calibration (library)`, that applies a linear excitation-energy calibration $E_i^{cal} = aE_i^{calc} + b$, an oscillator-strength scaling $f_{cal} = \alpha * f$, and optional per-band sigma / reorganisation-energy overrides before broadening; the identity calibration ($a=1$, $b=0$, $\alpha=1$) is bit-identical to running without it. The tabular branch additionally accepts any externally pre-corrected spectrum.

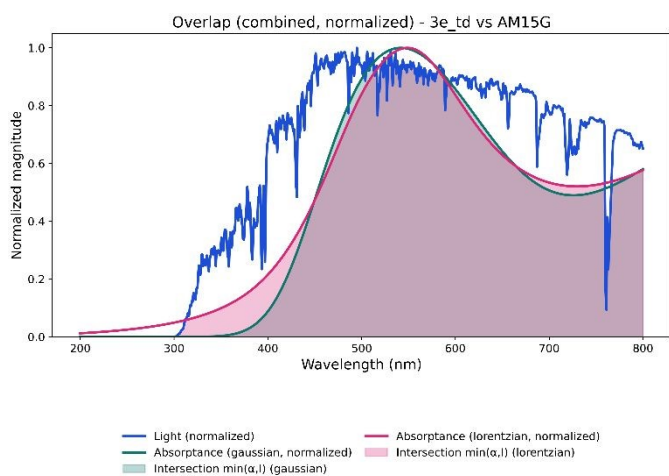


Fig. 3 Representative spectral-overlap plot between a TD-DFT-derived absorbance spectrum and AM1.5G illumination.

Spreadsheet-input demonstration using public UV-vis spectra

To demonstrate the spreadsheet-input branch of the workflow, a separate UV–vis dataset of representative organic chromophores was prepared from publicly available PhotochemCAD/OMLC absorption spectra associated with the Taniguchi and Lindsey PhotochemCAD database.⁵³

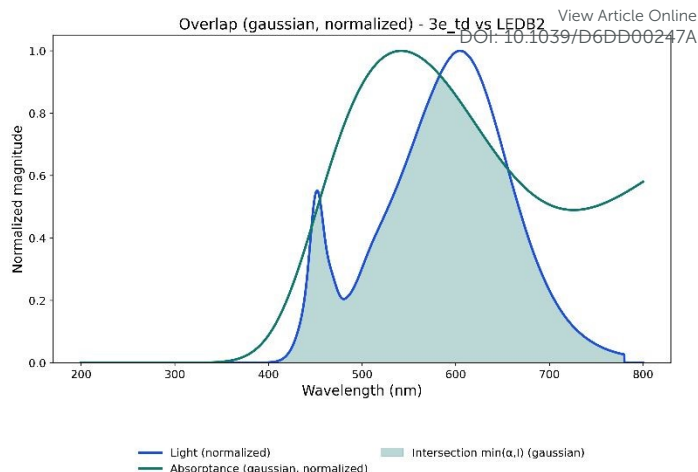


Fig. 4 Representative spectral-overlap plot between a TD-DFT-derived absorbance spectrum and LED B2 illumination.

The molar extinction coefficient spectra were converted into absorbance-like profiles using $A(\lambda) = \epsilon(\lambda)cL$ with a fixed reference concentration of 1×10^{-5} M and a 1 cm optical path length. The reference concentration c and path length L are state parameters of the Beer–Lambert conversion, not fit parameters. In the optically thin limit `shape_overlap` is invariant under any global scaling of ϵcL , and `absorbed_fraction` is a monotone-saturating function of cL that scales linearly with it for $A \ll 1$; only `absorbed_flux` carries the absolute scale. For screening we therefore recommend choosing cL so that the peak reference absorbance lies in the range $A_{max} \in [0.1, 1]$ — the well-conditioned regime of experimental cuvette spectroscopy. For typical organic chromophores ($\epsilon_{max} \sim 10^4$ – 10^5 M⁻¹cm⁻¹) the defaults $c = 1 \times 10^{-5}$ M and $L = 1$ cm satisfy this, and both values are written to every output row, so no iterative optimisation of c and L is required. The resulting values were arranged in an Excel workbook with `wavelength_nm` in the first column and each molecule as an independent spectral series.

This dataset is used only as a spreadsheet-input demonstration and is not intended as an experimental validation set for the TD-DFT-derived molecules (Fig. 5). It verifies that organic molecular UV-vis spectra supplied in tabular form can be ingested, converted to absorbance, compared with the same solar and indoor light-source spectra, and exported with the same output logic used for TD-DFT-derived spectra. The spreadsheet spectra were used only to demonstrate the tabular-input branch of the workflow and are not proposed as OPV candidate absorbers, as clarified in Section S4 of the ESI.

Unified descriptor generation from theoretical and spreadsheet inputs

Both input branches are routed to a common output structure. For TD-DFT files, the workflow first reconstructs Gaussian- and Lorentzian-broadened spectra from discrete excited states before applying the Beer-Lambert absorbance model. For spreadsheet spectra, the supplied profile is treated as



absorbance and converted directly into absorbance. After this point, both branches are compared with the same light-source library and exported as descriptor tables, summary files, skip reports, timing tables, and TIFF visualisations.

This design is useful because it allows a project to start from TD-DFT screening data and later incorporate measured UV-vis spectra without changing the downstream analysis. The same descriptors can therefore be used for exploratory ranking, experimental follow-up, and reproducibility checks, while clearly distinguishing between calculated spectra, measured absorbance spectra, and public spreadsheet demonstration data.

Source-dependent ranking of candidate absorbers

The ranking analysis shows that the spectral compatibility of absorbers depends on the selected illumination source. Under AM 1.5G, broad absorption across the visible region may increase absorbed flux. Under LED or fluorescent sources, however, a narrower but better-aligned absorbance profile may provide a higher shape_overlap (Fig. 6).

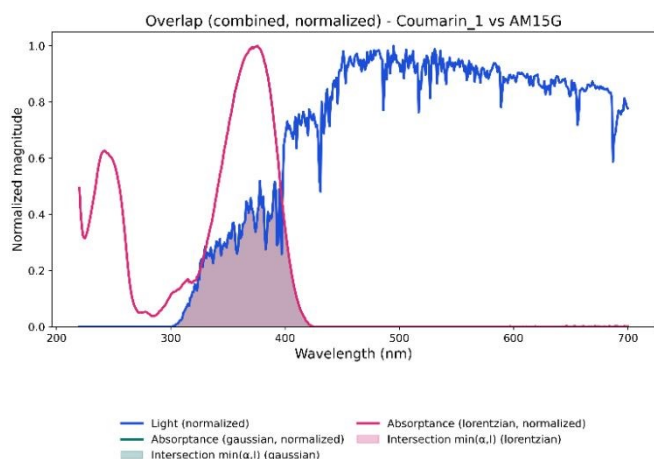


Fig. 5 Representative spectral-overlap plot between a spreadsheet-input absorbance spectrum derived from an experimentally reported UV-vis spectrum and AM1.5G illumination.

Therefore, the workflow supports a design logic in which absorbers are not ranked by a single universal optical descriptor but by their compatibility with the intended operating environment. The top-ranked samples under each light source are provided in Table S7 of the ESI.

The ranking analysis confirms that the spectral compatibility of the tested absorbers is strongly source-dependent. Under AM1.5G illumination, 4c_td gave the highest Gaussian shape-overlap value of 0.7952, followed by 3e_td with 0.7173 and 3g_td with 0.6437. In contrast, under LEDB4 illumination, 4c_td remained the top-ranked sample with a Gaussian shape-overlap value of 0.9643, followed by 3g_td with 0.9536 and 3e_td with 0.9374. For the representative TD-DFT-derived absorber 3e_td, the highest shape-overlap values were obtained under LEDB2, LEDB3, LEDB4, and CIEFL10, whereas the AM1.5G overlap was lower. This demonstrates that absorber ranking cannot be

inferred from λ_{max} alone and should be evaluated with respect to the intended illumination environment.

Scope and extensibility of the workflow

The main contribution of overlap-calculator is the standardisation of a recurring interpretation step in OPV screening. A small in-house script can answer a specific question for a single dataset, but it is difficult to reuse, document, and audit unless its assumptions are made explicit. By separating input definition, parsing, spectral reconstruction, light-source integration, descriptor calculation, and output export, the present workflow makes each assumption visible. This is particularly valuable when the analysis is used to support publication claims, compare molecule families, or share data with collaborators.

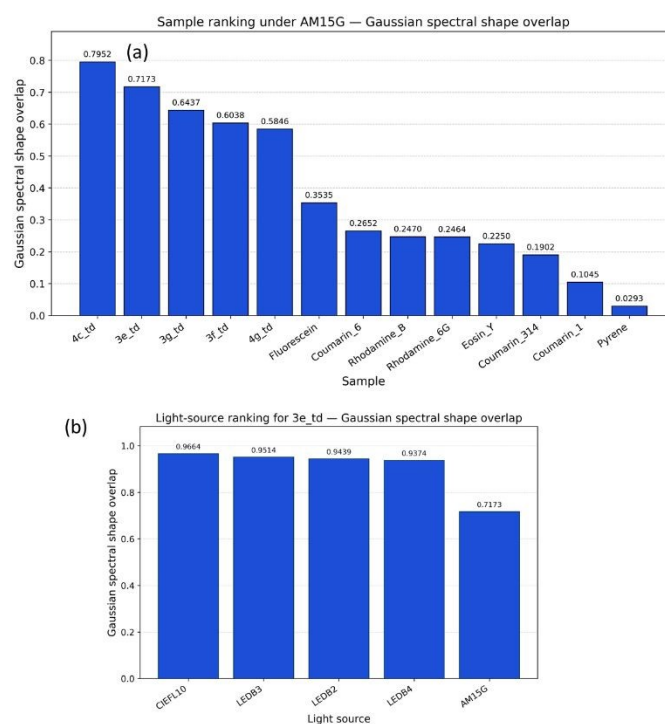


Fig. 6 Source-dependent spectral-overlap ranking generated by overlap-calculator. (a) Sample ranking under AM1.5G illumination based on Gaussian spectral shape-overlap values. (b) Light-source-dependent Gaussian shape-overlap values for the representative TD-DFT-derived absorber 3e_td.

For absolute irradiance spectra such as AM 1.5G, absorbed_flux has a direct physical meaning within the simplified absorbance model. For relative indoor illuminant spectra, however, the absolute scale is arbitrary. In those cases, absorbed_fraction and shape_overlap are more informative for comparing molecules under the same light source. Shape overlap is especially useful when the goal is to quantify how well the positions and widths of absorption bands align with the emission profile of a lamp. However, it should not be interpreted as a device efficiency or as a replacement for experimental external quantum efficiency.



The descriptors reported by overlap-calculator are optical compatibility descriptors. They do not include exciton diffusion, morphology, donor-acceptor interfacial energetics, charge separation, transport, recombination, electrode losses, or optical interference in a real device stack. Therefore, they should be used as early-stage screening descriptors rather than final performance predictors. Their strength is that they quantify a necessary optical condition: if a candidate absorber does not overlap with the intended illumination source, high device performance is unlikely regardless of favourable electronic descriptors.

The current workflow has several limitations. TD-DFT reconstruction depends on the selected functional, basis set, solvent model, number of excited states, wavelength range, and broadening width. Beer-Lambert absorbance uses a simplified reference concentration and path length, and film optical constants and device-stack effects are not included. Experimental spectra are assumed to be absorbance spectra with a valid wavelength axis, so additional preprocessing may be needed for noisy or baseline-shifted data. Relative CIE LED and fluorescent spectra do not provide absolute absorbed power unless a calibrated intensity scale is supplied. The current descriptors quantify spectral compatibility, not EQE, Jsc, PCE, or long-term device stability. Interpretive limitations of the optical-overlap descriptors are discussed in Section S14 of the ESI.

Conclusions

overlap-calculator provides a reproducible Python workflow for quantifying spectral compatibility between molecular absorbers and reference light sources. By accepting both Gaussian TD-DFT outputs and spreadsheet-based UV-vis spectra, reconstructing or importing absorption profiles, converting them into absorbance, and integrating them against AM 1.5G, CIE LED, fluorescent, or custom spectra, the workflow transforms a previously manual interpretation step into a documented and batch-compatible analysis.

The resulting descriptors, including absorbed flux, absorbed fraction, and shape overlap, support source-dependent ranking of TD-DFT-derived OPV candidate absorbers while preserving transparency about the assumptions used in the calculation. The package is intended as a lightweight post-processing layer for computational and experimental optoelectronic materials research rather than as a replacement for electronic-structure calculations, spectroscopic measurements, or device-level photovoltaic characterization.

Future versions can extend the workflow in several directions. Support for EQE or IPCE spectra would allow the same light-source integration logic to estimate source-specific current-density limits. A graphical interface could make the workflow accessible to users who are less comfortable with command-line tools. Additional lamp libraries, calibrated spectral irradiance data, and user-defined spectral normalisation modes could further improve indoor PV analysis, while direct export of ranked summary reports could make the package more useful for high-throughput screening studies.

Author contributions

Pinar Seyitdanlioglu: Conceptualization, methodology, software, validation, formal analysis, investigation, data curation, writing - original draft, writing - review and editing, visualization.

Conflicts of interest

There are no conflicts to declare.

Data availability

The source code, README file, user manual, worked case studies, example input files, spreadsheet demonstration data, light-source spectra, and output tables supporting this study are available at: <https://github.com/pinarsyt/overlap-calculator>. A versioned archival release containing the same reproducibility materials is available at Zenodo: <https://zenodo.org/records/20579269>.

Acknowledgements

The author thanks Gökay Öztürk for his valuable support and encouragement during the preparation of this manuscript. AI-assisted language-editing tools were used to improve grammar, clarity, and readability; all scientific content and interpretations were reviewed and approved by the author. The TD-DFT calculations used in this paper were performed at TUBITAK ULAKBIM, High Performance and Grid Computing Center (TRUBA resources).

References

1. B. Kippelen and J.-L. Brédas, *Energy Environ. Sci.*, 2009, **2**, 251–261.
2. G. J. Hedley, A. Ruseckas and I. D. W. Samuel, *Chem. Rev.*, 2017, **117**, 796–837.
3. G. Zhang, F. R. Lin, F. Qi, T. Heumüller, A. Distler, H.-J. Egelhaaf et al., *Chem. Rev.*, 2022, **122**, 14180–14274.
4. M. Jahandar, S. Kim and D. C. Lim, *ChemSusChem*, 2021, **14**, 3449–3474.
5. X. Liu, S. Xu, B. Tang and X. Song, *Chem. Eng. J.*, 2024, **497**, 154944.
6. Y. J. You, C. E. Song, Q. V. Hoang, Y. Kang, J. S. Goo, D. H. Ko et al., *Adv. Funct. Mater.*, 2019, **29**, 1901171.
7. S. Biswas, Y. Lee, H. Choi, H. W. Lee and H. Kim, *RSC Adv.*, 2023, **13**, 32000–32022.
8. A. Venkateswararao, J. K. Ho, S. K. So, S.-W. Liu and K.-T. Wong, *Mater. Sci. Eng. R Rep.*, 2020, **139**, 100517.
9. M. Freunek, M. Freunek and L.M. Reindl, *IEEE J. Photovolt.*, 2012, **3**, 59–64.
10. H.K.H. Lee, J. Wu, J. Barbé, S.M. Jain, S. Wood, E.M. Speller et al., *J. Mater. Chem. A.*, 2018, **6**, 5618–5626.
11. A. Venkateswararao, K. T. Wong, *Bull. Chem. Soc. Jpn.*, 2021, **94**, 812–838.
12. M. Li, F. Igbari, Z.K. Wang and L. S. Liao, *Adv. Energy Mater.*, 2020, **10**, 2000641.
13. Y. Cui, L. Hong and J. Hou, *ACS Appl. Mater. Interfaces*, 2020, **12**, 38815–38828.



14. Y. Cui, Y. Wang, J. Bergqvist, H. Yao, Y. Xu, B. Gao et al., *Nat. Energy*, 2019, **4**, 768–775.
15. A. Chakraborty, G. Lucarelli, J. Xu, Z. Skafi, S. Castro-Hermosa, A. B. Kaveramma et al., *Nano Energy*, 2024, **128**, 109932.
16. J. Hachmann, R. Olivares-Amaya, S. Atahan-Evrenk, C. Amador-Bedolla, R. S. Sánchez-Carrera, A. Gold-Parker et al., *J. Phys. Chem. Lett.*, 2011, **2**, 2241–2251.
17. S. A. Lopez, E. O. Pyzer-Knapp, G. N. Simm, T. Lutzow, K. Li, L. R. Seress et al., *Sci. Data*, 2016, **3**, 160086.
18. M. Bourass, A. T. Benjelloun, M. Benzakour, M. Mcharfi, M. Hamidi, S. M. Bouzzine and M. Bouachrine, *Chem. Cent. J.*, 2016, **10**, 67.
19. J. Yan, X. Rodríguez-Martínez, D. Pearce, H. Douglas, D. Bili, M. Azzouzi et al., *Energy Environ. Sci.*, 2022, **15**, 2958–2973.
20. S. Rani, N. Al-Zaqri, J. Iqbal, S. J. Akram, A. Boshala, R. F. Mehmood et al., *RSC Adv.*, 2022, **12**, 29300–29318.
21. S. R. Bora and D. J. Kalita, *RSC Adv.*, 2023, **13**, 26418–26429.
22. D. Lübke, P. Hartnagel, J. Angona and T. Kirchartz, *Adv. Energy Mater.*, 2021, **11**, 2101474.
23. M. A. Saeed, *Appl. Phys. Lett.*, 2025, **127**, 130501.
24. S. Y. Park, C. Labanti, J. Luke, Y. C. Chin and J. S. Kim, *Adv. Energy Mater.*, 2022, **12**, 2103237.
25. X. Hou, Y. Wang, H. K. H. Lee, R. Datt, N. U. Miano, D. Yan et al., *J. Mater. Chem. A*, 2020, **8**, 21503–21525.
26. A. S. Tarleton, J. C. Garcia-Alvarez, A. Wynn, C. M. Awbrey, T. P. Roberts and S. Gozem, *J. Phys. Chem. A*, 2022, **126**, 435–443.
27. C. Goble, S. Cohen-Boulakia, S. Soiland-Reyes, D. Garijo, Y. Gil, M. R. Crusoe et al., *Data Intelligence*, 2020, **2**, 108–121.
28. J. Popp and T. Biskup, *Chemistry–Methods*, 2022, **2**, e202100097.
29. M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman et al., Gaussian 16, Revision C.02, Gaussian, Inc., Wallingford CT, 2019.
30. G. Pizzi, A. Cepellotti, R. Sabatini, N. Marzari, and B. Kozinsky, *Comput. Mater. Sci.*, 2016, **111**, 218–230.
31. A. S. Rosen, M. Gallant, J. George, J. Riebesell, H. Sahasrabudde, J. X. Shen, et al., *J. Open Source Softw.*, 2024, **9**, 5995.
32. J. Janssen, S. Surendralal, Y. Lysogorskiy, M. Todorova, T. Hickel, R. Drautz and J. Neugebauer, *Comput. Mater. Sci.*, 2019, **163**, 24–36.
33. J. Janssen, J. George, J. Geiger, M. Bercx, X. Wang, C. Ertural et al., *Digit. Discov.*, 2025, **4**, 3149.
34. R. C. Hilborn, *Am. J. Phys.*, 1982, **50**, 982–986.
35. R. S. Mulliken, *J. Chem. Phys.*, 1939, **7**, 20–34.
36. International Commission on Illumination, Relative spectral power distributions of illuminants representing typical LED lamps, 1 nm spacing, CIE Central Bureau, Vienna, 2018, DOI: 10.25039/CIE.DS.dhcw57sd.
37. International Commission on Illumination, Relative spectral power distributions of illuminants representing typical fluorescent lamps, 1 nm wavelength steps, CIE Central Bureau, Vienna, 2018, DOI: 10.25039/CIE.DS.54hy6srn.
38. International Commission on Illumination, CIE 015:2018 Colorimetry, 4th edn, CIE Central Bureau, Vienna, 2018.
39. National Renewable Energy Laboratory, Reference Air Mass 1.5 Spectra (ASTM G-173-03), U.S. Department of Energy.
40. R. A. Marcus, N. Sutin, *Biochim. Biophys. Acta - Rev. Bioenerg.*, 1985, **811**, 265–322.
41. S. F. Nelsen, S. C. Blackstock, Y. Kim, *J. Am. Chem. Soc.*, 1987, **109**, 677–682.
42. F. Santoro, A. Lami, R. Improta, V. Barone, *J. Chem. Phys.*, 2007, **126**, 184102. View Article Online
DOI: 10.1039/D6DD00247A
43. J. Bloino, M. Biczysko, F. Santoro, V. Barone, *J. Chem. Theory Comput.*, 2010, **6**, 1256–1274.
44. J. R. Reimers, J. R. *J. Chem. Phys.*, 2001, **115**, 9103–9109.
45. M. Barker, N. P. Chue Hong, D. S. Katz, A. L. Lamprecht, C. Martínez-Ortiz, F. Psomopoulos, et al., *Sci. Data*, 2022, **9**, 622.
46. M. D. Wilkinson, M. Dumontier, I. J. Aalbersberg, G. Appleton, M. Axton, A. Baak et al., *Sci. Data*, 2016, **3**, 160018.
47. S. Soiland-Reyes, P. Sefton, M. Crosas, L. J. Castro, F. Coppens, J. M. Fernández et al., *Data Sci.*, 2022, **5**, 97–138.
48. P. Seyitdanlioglu, *J. Mol. Model.*, 2026, **32**, 8.
49. A. D. Laurent, D. Jacquemin, *Int. J. Quantum Chem.*, 2013, **113**, 2019–2039.
50. C. Adamo, D. Jacquemin, *Chem. Soc. Rev.*, 2013, **42**, 845–856.
51. R. Send, M. Kühn, F. Furche *J. Chem. Theory Comput.*, 2011, **7**, 2376–2386.
52. D. Jacquemin, V. Wathelet, E. A. Perpète, C. Adamo, *J. Chem. Theory Comput.*, 2009, **5**, 2420–2435.
53. M. Taniguchi and J. S. Lindsey, *Photochem. Photobiol.*, 2018, **94**, 290–327.



Data availability statement

The source code, README file, user manual, worked case studies, example input files, spreadsheet demonstration data, light-source spectra, and output tables supporting this study are available at: <https://github.com/pinarsyt/overlap-calculator>. A versioned archival release containing the same reproducibility materials is available at Zenodo: <https://zenodo.org/records/20579269>.

