Telling the story of solar energy meteorology into the satellite era by applying (co-citation) reference publication year spectroscopy

Thomas Scheidsteger[1] · **Robin Haunschild**[1]

Abstract

Studying the history of research fields by analyzing publication records and topical and/or keyword searches with reference publication year spectroscopy (RPYS) has been introduced as a powerful tool to identify the corresponding root publications. However, for some research fields (e.g., rather new and interdisciplinary fields) like solar energy meteorology, encompassing such research fields via a keyword- or topic-based search query is not feasible to get a reasonably exhaustive publication set. Therefore, we apply its variant RPYS-CO to all publications co-cited with two highly important marker papers, using the cited references explorer for inspecting the RPYS-CO results. We obtain two lists of seminal papers, which are able to adequately tell us the story of solar energy meteorology up to the 1990s, respectively in its subfield using satellite-based methods for solar irradiance estimation even to very recent years. Consequently, we recommend this method to gain valuable insights in (new) research fields.

Introduction

Solar Energy Meteorology (SEM) studies how solar radiation can be utilized for solar energy conversion to provide heat or electricity to energy systems and how the performance of these conversion processes is affected by meteorological influences. The mainly interesting properties of solar radiation in this respect are its availability in time (e.g., time of day, year) and space (e.g., geographical location, angular orientation). Especially the latter is decisive for its exploitation capacity in different devices of energy conversion: Concentrating devices like parabolic trough power plants need

---

✉ Thomas Scheidsteger
t.scheidsteger@fkf.mpg.de

Robin Haunschild
r.haunschild@fkf.mpg.de

[1]   Max Planck Institute for Solid State Research, Heisenbergstr. 1, 70659 Stuttgart, Germany

direct radiation to work whereas non-concentrating devices as photovoltaic solar cells or flat-plate solar thermal collectors can also utilize the *diffuse* fraction of sunlight, usually scattered in the atmosphere. The total sum of both parts is called global irradiance. So, major fields of investigation in SEM are (1) measurements and their evaluation over different time scales and (2) modeling of radiation and its components, depending on physical (e.g., available solar radiation due to Planck's law and extinction processes in the atmosphere), geometrical (e.g., position of the sun, orientation of the converter) and meteorological (e.g., cloud coverage, aerosol concentration) parameters. Both fields also involve a large amount of statistical treatment.

The main instruments of solar radiation measurement had for a long time been ground-based pyranometers, integrating the incident sunlight from the hemisphere above a planar surface to the *global irradiance*, or pyrheliometers, detecting the direct sunlight within a narrow solid angle centered around the source of radiation—the *direct beam irradiance*. The instruments as well as their maintenance are costly and time consuming, and therefore high-quality data were only available for a limited number of locations. With the availability of geostationary weather satellites, for example the GOES generation, operated by the US American NOAA from 1975 onwards (https://www.goes-r.gov), and the METEOSAT series, operated by EUMETSAT from 1977 onwards (https://www.eumetsat.int/), solar radiation data in the visible and infrared part of the electromagnetic spectrum in a high temporal resolution (15 to 30 min) and with an almost continuous spatial coverage became accessible. Now the main task was the assessment of the radiation reaching the earth's surface from the on-board radiometer's digital counts, i.e. the digitally coded values of the measured radiation, always including a certain offset.

Two main methodologies to tackle this task have been developed from the late 1970s on, usually subdivided in the *statistical* and the *physical* approach. Statistical models perform a regression between the radiometer counts for a target area on earth and the available ground measurements, which are still today seen as "ground truth". The physical approach is largely based on radiative transfer models, which explicitly describe the extinction processes of scattering and absorption in the atmosphere and on the ground and need no ground measurements, but complementary meteorological data to estimate the relevant quantities. Physical models are more general and need no adaptation to target areas as the statistical ones do, but the latter are—due to their simplicity—more suited for time-critical, practical calculations and therefore extensively used in near real-time operational services and the building up of databases of long-term time series of a large spatial coverage.

One of the authors was involved in this work as a member of Oldenburg University's research group on Energy Meteorology for some time, so it came naturally to investigate this field of research with bibliometric methods in order to identify seminal and landmark papers, that lead up to and through the satellite era in SEM. But due to its intrinsic interdisciplinarity, potential search terms tend to have multiple meanings, which leads to answer sets from title or topical searches with low precision (i.e., fraction of retrieved relevant papers out of all of the retrieved papers) and/or recall (i.e., fraction of the retrieved relevant papers out of the total number of relevant papers). To illustrate this, we compare search results for the two subjects "Climate Change" and "Energy Meteorology" in the WoS, having searched (on 9 April 2020) in title, abstract and keywords. With ts = "climate change" in the WoS advanced search mode more than 200,000 papers which definitely deal with the very subject of climate change are found. Approximately half of the research area is covered by this simple WoS search (Haunschild et al. 2016b, 2019c). However, ts = "energy meteorology" yields only 36

papers with absolute thematic precision but very low recall. A broader search with proximity operators would significantly decrease the precision, but not sufficiently increase the recall.

Therefore, we use a certain variant of the bibliometric method "Reference Publications Year Spectroscopy" (RPYS), where there is no need for an exhaustive paper set covering most of the research field. RPYS had been introduced by Marx et al. (2014) in order to identify the historical roots of research fields. To this purpose the citation frequency of references in the publication corpus of a specific research field is analyzed in terms of the publication years of these cited references. A plot of the reference publication years (RPY) versus the number of cited references (NCR) usually shows peaks. The more or less pronounced peaks will be primarily caused by important papers, which stand out against others from the same year by a relatively high citation count inside the whole of the analyzed publication corpus. Previously, RPYS analyses were applied to a scientific field (e.g., Higgs boson research (Barth et al. 2014), dentistry and neurosciences (Yeung 2017; Yeung et al. 2019), health equity (Yao et al. 2019), or density functional theory (Haunschild et al. 2016a, 2019a; Haunschild and Marx 2019), a journal (e.g., the journal *Ecological Economics* (Ballandonne 2019), the journal *FEMS Microbiology Letters* (Haunschild et al. 2019b), *and the journal MDPI Information (Haunschild* 2019)), or the oeuvre of a researcher (e.g., oeuvre of Eugene Garfield (Bornmann et al. 2018)).

RPYS has also successfully been applied to identify the root publications of climate change research by Marx et al. (2017). In that case, a set of more than 200,000 papers has been used. In a subsequent approach in the same paper, Marx et al. (2017) refined this large set to the greenhouse effect by keeping only cited references that are co-cited with Arrhenius (1896) and were able to retrieve the results of the RPYS on the full publication set regarding the greenhouse effect, but also lesser known works of relevance. They named this RPYS variant RPYS-CO (for co-citation), because here the analyzed publication set is defined by all publications co-cited with at least one marker paper. In another very recent study, Haunschild and Marx (2019, 2020) compared their own results of an RPYS on density functional theory, a very frequently applied method in computational chemistry (Haunschild et al. 2016a), with an RPYS-CO using *one single* seminal paper with a high citation count and strong relevance in the field as a marker paper. They found a striking similarity with the results of the analysis based on a search in a controlled vocabulary.

Encouraged by these results, we set out here to investigate all publications co-cited with one highly-cited marker paper, a choice discussed with and corroborated by the long term leader of the Oldenburg group: *The interrelationship and characteristic distribution of direct, diffuse and total solar radiation* (Liu and Jordan 1960). In order to indicate the importance of this marker paper we quote its complete abstract, emphasizing in italics all those concepts and terms that proved to be prevalent in SEM for its whole history:

"Based upon the data now available, this paper presents relationships permitting the determination on a *horizontal surface* of the instantaneous intensity of *diffuse radiation* on *clear days*, the *long term average hourly and daily sums* of diffuse radiation, and the daily sums of diffuse radiation for various categories of days of *differing degrees of cloudiness*. For these determinations, it is necessary to have, either from actual measurements or estimates, a knowledge of the *total (direct plus diffuse) radiation on a horizontal surface*—its *measurement is now regularly made* at 98 localities in the United States and Canada. For localities where only an *estimate of the long term average total radiation* is available, *relationships* presented in this paper can be utilized to determine the *statistical distribution of the daily total radiation at these localities*." (Liu and Jordan 1960, abstract).

Satellite-based studies often view this paper as text book knowledge. Therefore, it is affected by obliteration by incorporation (McCain 2014) and rarely cited in this area of SEM, that gained traction in the 1980s and early 1990s. The latter are consequently a natural end date for our study of cited references, co–cited with Liu and Jordan (1960). On the other hand, due to the recency of the research field, we do not expect decisive contributions leading to SEM before 1900. The more recently flourishing satellite-based publications need to be studied using another marker paper, for which an appropriate choice is: *A method for the determination of the global solar radiation from meteorological satellite data* (Cano et al. 1986). This paper had also been recommended by the long-term expert being a reference point for all the work of his group and other groups in the world-wide community.

There are some studies with a very different time frame, focus or methodology, e.g.: Du et al. (2014) analyzed the solar energy literature from 1992 to 2011, but without special consideration of energy meteorology topics. A bibliometric analysis on solar power research between 1991 and 2010, again after the period of the first part of our study, has been performed by Dong et al. (2012) using terms as, e.g., "solar radiation" in a topic search in the WoS. Their goal was to identify research trends for the twenty-first century and not to explore historical roots. In the same vein, Yang et al. (2018) tried to identify key innovations for the future of research in "solar radiation and PV power forecasting", a field mainly emerging at the turn of the millennium. They based their work on the first 1000 hits of a keyword search in Google Scholar and applied machine learning and text mining methods to full texts in order to complement conventional topical reviews.

To the best of our knowledge, the present study is one of the first using the method RPYS-CO in order to identify seminal papers for a research field—thereby complementing qualitative knowledge of experts by a quantitative evaluation of the citation counts (i.e., the reference counts within the topic-related literature). Using this method, we are confident to find those important contributors and their papers which tell the story of the emergence of solar energy meteorology from around 1900 up to the beginning of the 1990s and how the availability of huge amounts of satellite-derived data was taken up by the SEM community by developing new scientific calculation schemes and operational services driven by research groups as well as public and private institutions. So, we support the suggestion of Haunschild and Marx (2019) that this method can help researchers to explore their field of study—in a way complementary to a usual topic or keyword-based literature search.

The present study is an extended version of a paper presented at the 17th International Conference on Scientometrics and Informetrics, Rome (Italy), 2–5 September 2019 (Scheidsteger and Haunschild 2019). The new material comprises the second focus on satellite-based methods, including the related explanations in the "Introduction", the bibliometric characterization of the second marker paper in "Method and data set" and its investigation and presentation as sub-section "RPYS-CO using the second marker paper" in the "Results" section, including Fig. 5; Table 2, and parts of the "Discussion" section. Moreover Fig. 6 displays a further confirmation of the results in the sub-section "RPYS-CO using the first marker paper".

Method and data set

As of 8 January 2019, the marker paper Liu and Jordan (1960) had 1031 citing papers in the WoS until the end of 2018. One fourth of these papers (n = 257, 25%) as well as the marker paper itself have been published in a single journal, *Solar Energy*. The four

most important WoS subject categories in the data set used in this study are *Energy Fuels* (n=673, 65%), *Green Sustainable Science Technology* (n=151, 15%), *Meteorology and Atmospheric Sciences* (n=131, 13%), and *Thermodynamics* (n=114, 11%), thereby reflecting the multiple foci of SEM.

We downloaded the bibliographic data of the 1032 papers including 36,635 cited references (CRs) from the WoS (selecting "Save to Other File Formats" and "Other Reference Software") and imported them into the CRExplorer. (The Java-based software can be downloaded for free from http://crexplorer.net and a comprehensive handbook explaining all functions is also available.) It provides a graphical display of the NCR over the RPY and a tabular presentation of the NCR of all CRs. In our case there were only single occurrences of CRs before 1900.

Much of the processing can comprehensively and reproducibly be done by using the CRExplorer scripting language: With the script in Fig. 1 we imported the WoS file and got 8383 unique reference variants for the reference publication years 1900 to 1995. After that clustering and merging of equivalent CR variants was done with Levenshtein threshold 0.75 and taking volume and (starting) page number into account, thereby reducing the number of CR variants by 109. Then we removed all publications with only *one* citation, in order to reduce noise. In the end, we retained 1566 CRs. The results including the NCR and other indicators were exported to CSV files for further inspection and plotting of the spectrogram, which can be done by using the R package BibPlots (see: https://cran.r-project.org/web/packages/BibPlots/index.html and https://tinyurl.com/y97bb54z).

In the spectrogram, we looked for publication years with significantly higher NCR than other years, aided by the deviation of NCR from the 5-year-median of NCR (taking into account the two preceding and the two following years). For the papers that, by applying this methodology, seemed primarily responsible for the peaks a manual merging was done, if needed.

For the subfield of satellite-based SEM we took Cano et al. (1986) as marker paper.

As of 30 January 2019, it had 293 citing papers in the WoS until the end of 2018, which seems to be rather low compared to the choice of other marker papers in the literature, but taking into account that it only represents a small substream of SEM, which we are going to treat with just over 1000 citing papers, we expect no serious disadvantage.

One fifth of the papers citing Cano et al. (1986) (n=60, 20%) as well as this marker paper itself have been published in the journal *Solar Energy*. The three most important WoS subject categories in the data set used in the second part of this study are *Energy Fuels* (n=126, 43%), *Remote Sensing* (n=78, 27%), and *Meteorology and Atmospheric Sciences* (n=58, 20%), covering 87% of all publications.

```
importFile(file:"savedrecs_Liu_1960.txt",type:"WOS",
RPY:[1900,1995,false],PY:[1962,2018,false],maxCR:0)
info()
cluster(threshold:0.75,volume:true,page:true,DOI:false)
merge()
info()
removeCR(N_CR:[0,1])
info()
saveFile(file:"Liu1960.rpys.cre")
exportFile(file:"Liu_1960.rpys_CR.csv",type:"CSV_CR")
exportFile(file:"Liu_1960.rpys_GRAPH.csv",type:"CSV_GRAPH")
```

Fig. 1 CRExplorer script to perform RPYS on the WoS papers citing Liu and Jordan (1960)

We downloaded the bibliographic data of the 293 papers citing Cano et al. (1986) including 11,549 CR variants coming from 6723 unique CRs from the WoS, imported them into the CRExplorer and treated them analogously to the marker paper for the entire SEM field. We retained 1350 unique CRs between 1922 and 2017, cited at least twice.

Results

### RPYS-CO using the first marker paper

Figure 2 shows the spectrogram of the RPYS-CO for the SEM marker paper (Liu and Jordan 1960) in terms of the NCR and their 5-year-median deviation for the whole analyzed time period, and Table 1 lists all publications, contributing substantially to the peaks of NCR and identified as relevant.

The overall RPYS-CO picture can easily be divided by the maximum NCR per RPY into two periods with regard to the reference publication years, which we are going to discuss separately: the first one from 1915 leading to, but excluding, 1960, the publication year of the marker paper, containing peaks with at most NCR = 100; the second one from 1960 to 1995, with peaks between NCR = 100 and NCR = 900 (apart from the marker paper itself).

Time period 1915 to 1959

Figure 3 shows the RPYS-CO spectrogram for the marker paper Liu and Jordan (1960) for the relevant time period before it was published.

In this time period, we were able to identify 9 peaks with *relevant* papers for the following RPYs: 1919, 1922, 1924, 1929, 1940, 1942, 1945/46, 1948, and 1953-58. Because of the generally low NCR values in this time period, we did not want to lose reference
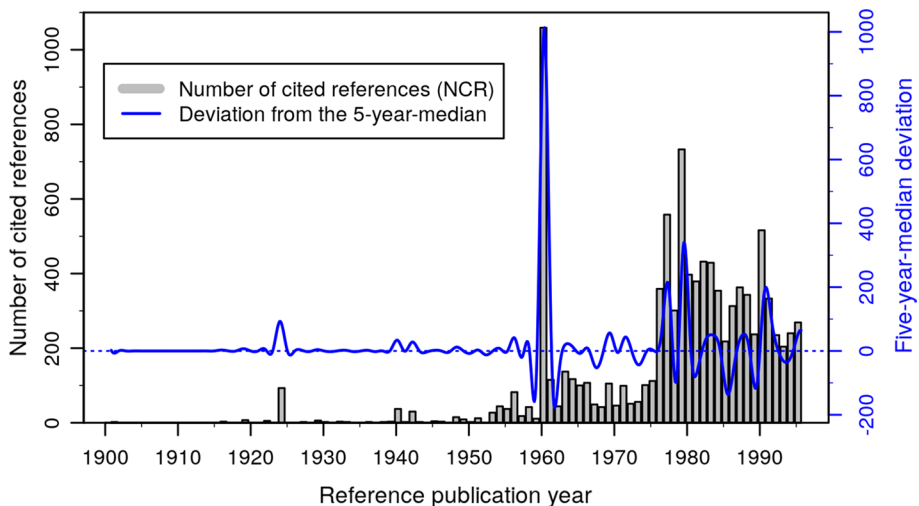
Fig. 2 Overall RPYS-CO graph for Liu and Jordan (1960) with NCR (grey bars) and 5-year-median deviation (blue line). (Color figure online)

Table 1 RPYS-CO for Liu and Jordan (1960): important CRs from 1900 to 1995 with number of citations NCR and indicating, if manually merged during inspection of spectrogram

#CR	RPY	Cited Reference (Manually merged: M)	NCR
CR1	1919	Kimball HH, 1919, Monthly Weather Review, V47, P769	7
CR2	1922	Angström A, 1922, Ark Mat Astron Fys, V17, P1	3
CR3	1922	Linke F, 1922, Beitr Phys Atmos, V10, P91	2
CR4	1924	Angström A, 1924, Quarterly Journal of the Royal Meteorological Society, V50, P121 *(M)*	93
CR5	1929	Angström A, 1929, Geogr Annlr Stockholm, V11, P156 *(M)*	6
CR6	1940	Prescott J, 1940, T Roy Soc South Aust, V64, P114	30
CR7	1942	Hottel HC, 1942, Transactions of the ASME, V64, P91 *(M)*	23
CR8	1945	Haurwitz B, 1945, J Met, V2, P154	4
CR9	1946	Haurwitz B, 1946, J Met, V3, P123	3
CR10	1948	Haurwitz B, 1948, J Met, V5, P110	5
CR11	1953	Whillier A, 1953, Thesis, MIT Cambridge *(M)*	17
CR12	1954	Black JN, 1954, Q J Roy Meteor Soc, V80, P231	28
CR13	1955	Hottel HC, 1955, T C Use Solar Energy, V2, P74	27
CR14	1956	Whillier A, 1956, Arch Meteorol Geophys U Bioklimatol Ser B, V7, P197 *(M)*	38
CR15	1958	Glover J, 1958, Q J Roy Meteor Soc, V84, P172	18
CR16	**1960**	**Liu BYH, 1960, Solar Energy, V4, P1**	**1031**
CR17	1963	Liu BYH, 1963, Solar Energy, V7, P53	71
CR18	1963	Choudhury NKD, 1963, Solar Energy, V7, P44	37
CR19	1964	Page JK, 1964, P UN C New Sources E, V4, P378 *(M)*	87
CR20	1966	Stanhill G, 1966, Solar Energy, V10, P96	29
CR21	1966	Robinson N, 1966, Solar Radiation *(M)*	25
CR22	1966	Kasten F, 1966, Arch Meteorol Geop B, VB14, P206	11
CR23	1969	Cooper PI, 1969, Solar Energy, V12, P333 *(M)*	42
CR24	1969	Kondratyev KY, 1969, Radiation in the Atmosphere	23
CR25	1971	Spencer J, 1971, Search, V2, P172 *(M)*	35
CR26	1974	Duffie JA, 1974, Solar Energy Thermal Processes *(M)*	35
CR27	1976	Ruth DW, 1976, Solar Energy, V18, P153	68
CR28	1976	Hottel HC, 1976, Solar Energy, V18, P129	63
CR29	1976	Tuller SE, 1976, Solar Energy, V18, P259	46
CR30	1976	Hay JE, 1976, Atmosphere, V14, P278	34
CR31	1977	Orgill JF, 1977, Solar Energy, V19, P357	149
CR32	1977	Klein SA, 1977, Solar Energy, V19, P325	121
CR33	1977	Temps RC, 1977, Solar Energy, V19, P179 *(M)*	46
CR34	1979	Collares-Pereira M, 1979, Solar Energy, V22, P155	238
CR35	1979	Klucher TM, 1979, Solar Energy, V23, P111	60
CR36	1979	Hay JE, 1979, Solar Energy, V23, P301	59
CR37	1980	Duffie JA, 1980, Solar engineering of thermal processes, 1st Ed.	63
CR38	1980	Iqbal M, 1980, Solar Energy, V24, P491	46
CR39	1981	Bendt P, 1981, Solar Energy, V27, P1	66
CR40	1982	Erbs DG, 1982, Solar Energy, V28, P293	208
CR41	1983	Iqbal M, 1983, Introduction to Solar Radiation *(M)*	117
CR42	1987	Skartveit A, 1987, Solar Energy, V38, P271	50

Table 1 (continued)

#CR	RPY	Cited Reference (Manually merged: M)	NCR
CR43	1987	Perez R, 1987, Solar Energy, V39, P221	42
CR44	1988	Suehrcke H, 1988, Solar Energy, V40, P413	36
CR45	1988	Graham VA, 1988, Solar Energy, V40, P83 (*M*)	28
CR46	1990	Reindl DT, 1990, Solar Energy, V45, P1	116
CR47	1990	Reindl DT, 1990, Solar Energy, V45, P9	54
CR48	1990	Perez R, 1990, Solar Energy, V44, P271	49
CR49	1991	Duffie JA, 1991, Solar engineering of thermal processes, 2nd Ed. (*M*)	83

The marker paper is printed in bold type
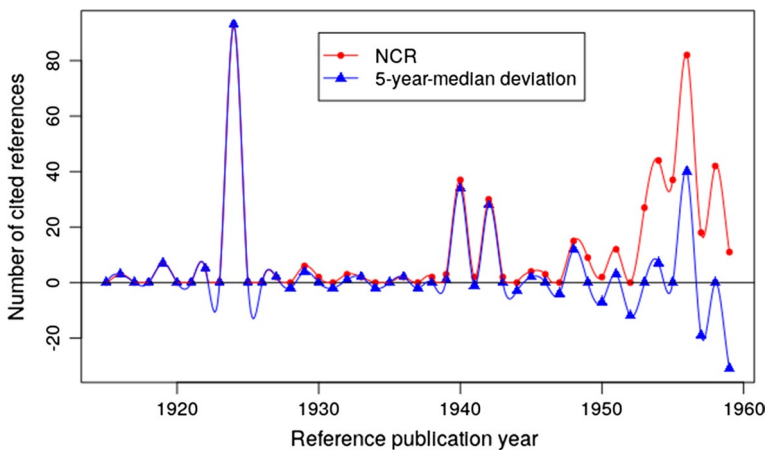
Fig. 3 RPYS-CO graph for Liu and Jordan (1960) and period 1 (1915–1959) with NCR (red line with circles) and 5-year-median deviation (blue line with triangles). (Color figure online)

variants of possibly relevant papers and therefore additionally looked at the CRs *before* the removal of only once referenced papers. But this did not reveal new relevant papers, instead it only confirmed the results from the reduced set.

Now we can follow the path of SEM by looking at the peak papers and drawing partially from explanations given in the citing papers. In this first period, two independent streams of research flew together: *meteorology* and *engineering*.

Meteorologists tried to develop a climatology of irradiance, emphasizing daily mean values, with no or little application to solar energy in mind. Solar irradiance varies deterministically with the sun's position on the sky dome (hemisphere) and irregularly with changing cloudiness. The relation of the latter with sunshine has initially been measured by Kimball (CR1) and later subjected to statistical analysis by Angström (CR2, CR4, and CR5), leading to a linear relation between the duration of bright sunshine and average solar energy available on a horizontal surface at ground level, the so-called *Angström equation*. This has been generalized to the *Angström–Prescott* (CR6) *equation* by introducing a *daily clearness index*, quantifying all stochastic meteorological influences, as a measure of the atmospheric transparency (Paulescu et al. 2016). Linke (CR3) in 1922 published his *turbidity factor* for the attenuation of the sun's radiation by water vapor and aerosols in the

atmosphere. Later Black, Bonython and Prescott (CR12) gave a linear regression relation between solar radiation and the duration of sunshine based on monthly values (Munkham-mar and Widen 2016) and Glover & McCulloch (CR15) improved this by including lati-tude effects.

Two *engineers*, Hottel and Woertz (CR7), came up with the first serious study on solar energy in 1942: the fundamental relationships given in their classic paper have since then been used for decades to model solar collectors. Hottel and Whillier (CR13) evaluated them concerning the flat-plate solar collector performance (Stanciu et al. 2016) and formu-lated the *Hottel-Whillier-Bliss equation* on its heat flow and available heat balance, based on considerations of the thermal usability of solar irradiation, coming from Whillier's PhD thesis (CR11) under Hottel's supervision at MIT. This latter work concerned also the rela-tion between radiation on different time scales, showing a close interdependence of the fre-quency distributions of the so-called clearness index on a monthly, daily, and hourly basis (Vijayakumar et al. 2005). Because information on daily sunshine duration was no longer sufficient, he later proceeded to "The determination of hourly values of total solar radiation from daily summations" (CR14) by statistical means, a subject still of great importance for SEM, where still an ever more time-resolved knowledge of solar irradiance is needed.

The *modeling* of solar irradiance components through parameterization of atmospheric phenomena is an equally important area of work in SEM. It was begun in the 1940s by Haurwitz (CR8-CR10), focusing on cloudiness, cloud density, and cloud type (Chowdhury 1990). (We do not include a publication from the 1948 peak by Penman with NCR = 8, i.e. more than CR10, because it is only concerned with evaporation by solar radiation.)

Time period 1960 to 1995

Figure 4 shows the RPYS-CO spectrogram for the marker paper (Liu and Jordan 1960) after its publication year.
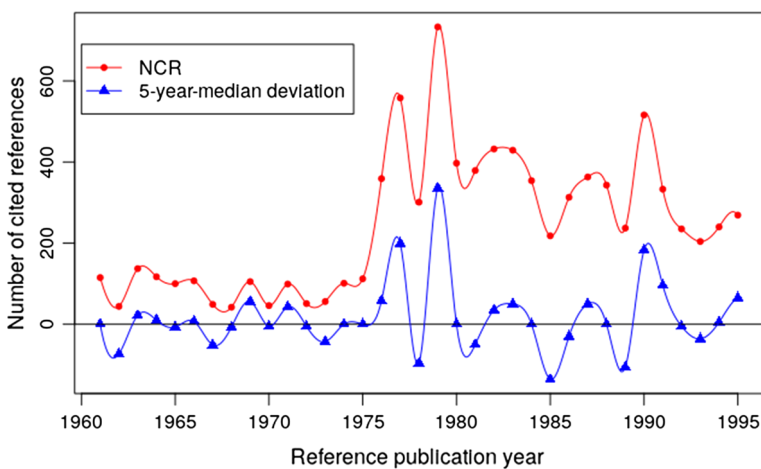
Fig. 4 RPYS-CO graph for (Liu and Jordan 1960) and period 2 (1960–1995) with NCR (red line with circles) and 5-year-median deviation (blue line with triangles). For better presentation, the RPY 1960 is excluded. (Color figure online)

Further 10 peaks could be identified from Fig. 4 for the following RPYs in the second time period: 1960, 1963/1964, 1966, 1969, 1971, 1976/1977, 1979/1980, 1981–1983, 1986–1988, 1990/1991.

The first two peaks are mainly caused by *engineers*: After the marker paper itself (CR16) the same authors gave generalized curves to predict the "Long-Term Average Performance of Flat-Plate Solar-Energy Collectors", making use of Hottel's and Whillier's methods (CR13) and building upon the knowledge of two parameters only: (1) the monthly-average daily total radiation on a horizontal surface and (2) the monthly average day-time ambient temperature (CR17). In 1961, J. K. Page presented "The estimation of monthly mean values of daily total short-wave radiation on vertical and inclined surfaces from sunshine records for latitudes 40 N-40 S" (CR19) on a "UN Conference on New Sources of Energy" in Rome but the conference proceeding was published in 1964. Much later, he advised advanced publicly funded projects like HELIOSAT-3 (Mueller et al. 2004).

The *meteorologist* F. Kasten (CR22) developed turbidity models as one essential ingredient for radiation model calculations and also functioned as an advisor in later solar energy projects.

Attempts to check and confirm the *diffuse-to-total radiation correlation* by Liu and Jordan (1960) against *measurements* have been undertaken for several locations in the world: Southern Israel by Stanhill (CR20), New Delhi by Choudhury (CR18), and Canada by Ruth and Chant (CR27) and Tuller (CR29).

Cooper (CR23) and Spencer (CR25) are the only representatives in Table 1 of those researchers concerned with *solar geometry*, i.e. sun-earth angle values over time—which is essential for all modeling of radiation. In this respect, our method could only capture these first works, but not later standard works like Michalsky (1988).

Robinson's "Solar Radiation" (CR21) was a meteorological standard publication, but not that much focused as Kondratyev's monograph "Radiation in the Atmosphere" (CR24), whose influence lasted until Iqbal's standard work "Introduction to Solar Radiation" (CR41) from 1983.

In the 1970s and early 80s, one focus of research literature was on *time-resolved diffuse radiation models from the scale of months down to hours*, mostly from the viewpoint of *engineering* like Duffie and Beckman's volumes "Solar Energy Thermal Processes" in 1974 (CR26) and "Solar engineering of thermal processes" in two editions in 1980 (CR37) and 1991 (CR49), but also Hottel (CR28), Orgill and Hollands (CR31), Klein (CR32), and Erbs, Klein and Duffie (CR40). Only Hay (CR30) and Iqbal (CR38) represent *geography* resp. *meteorology*. *Empirical radiation modeling*, in particular with respect to tilted surfaces (e.g., of solar panels), was done by Hay (CR36) and *meteorologists* as Temps & Coulson (CR33) and Klucher (CR35).

At the end of the 1970s, the focus switched also to *stochastic modeling*, outstandingly covered by Collares-Pereira and Rabl (CR34) with their time series analysis and production of the *first synthetic time series*, that were widely used in the community. In the same vein, Bendt presented his "frequency distribution of daily insolation values" (CR39). The timescale was later even narrowed down to *minute data* by Suehrcke and McCormick (CR44), and Graham, Hollands & Unny (CR45) were able to *simulate daily values* of the clearness index from monthly mean values by using probability distribution functions.

In 1987, Skartveit and Olseth (CR42) presented a *diffuse fraction model*, that became essential part of later works, e.g. in HELIOSAT (Dürr and Zelenka 2009). CR43, i.e., Perez et al. (1987), also focused on the diffuse part of total irradiance and accomplished a major improvement in its error-prone computation, in order to "estimate short time step (hourly or less) irradiance on tilted planes" (Perez et al. 1987), which has received

high recognition in the community. (See *Discussion and Conclusion* for considerations to use CR43 as a second marker paper.)

Duffie and Beckman, together with their coauthor Reindl, were mainly responsible for the last high peak, taken into account in our RPYS-CO analysis, in 1990: they *evaluated statistical models for hourly radiation on the tilted surface* (CR47) and could significantly *improve on the time resolution of statistical diffuse radiation models* in CR46, whose abstract we now quote (with our emphases in italics): "The influence of climatic and geometric variables on the *hourly diffuse fraction* has been studied, based on a data set with *22,000 hourly measurements* from five European and North American locations. The goal is to determine if other predictor variables, in addition to the clearness index, *will significantly reduce the standard error* of *Liu- and Jordan-type correlations* (…). Stepwise regression is used to *reduce a set of 28 potential predictor variables* down to *four significant predictors: the clearness index, solar altitude, ambient temperature, and relative humidity*." (Reindl et al. 1990, abstract).

We can in a sense close the circle to our first marker paper after exactly 30 years by mentioning CR48, i.e., Perez et al. (1990), as a successful attempt to *apply diffuse fraction models* to questions of *daylighting in buildings*, transferring irradiance to illumination, and again connecting the fields of *meteorology and radiation* to *energy and engineering*, the twofold focus of SEM.

RPYS-CO using the second marker paper

Figure 5 shows the spectrogram of the RPYS-CO for the second marker paper Cano et al. (1986) in terms of the NCR and their 5-year-median deviation for the whole analyzed time period, and Table 2 lists all publications contributing substantially to the peaks of NCR and identified as relevant, which had not already been listed in Table 1.
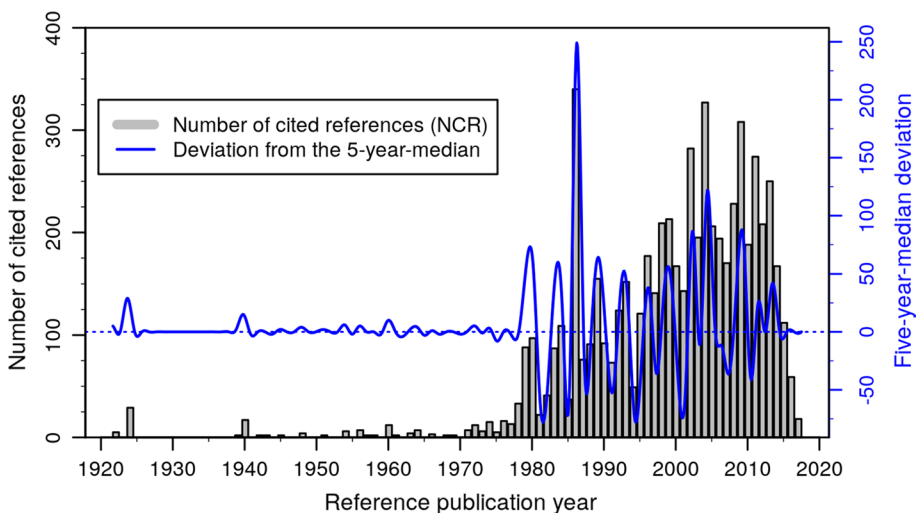
Fig. 5 Overall RPYS-CO graph for Cano et al. (1986) with NCR (grey bars) and 5-year-median deviation (blue line). (Color figure online)

Table 2 RPYS-CO for the second marker paper Cano et al. (1986): important CRs from 1964 to 2017 with number of citations NCR

#CR	RPY	Cited Reference	NCR
CR101	1964	Fritz S, 1964, J Atmos Sci, V21, P141	7
CR102	1972	Priestley CHB, 1972, Mon Weather Rev, V100, P81	5
CR103	1974	Lacis AA, 1974, J Atmos Sci, V31, P118	9
CR104	1976	Paltridge G.W., 1976, Rad Processes Meteor	4
CR105	1978	Hay JE, 1978, B Am Meteorol Soc, V59, P1549	10
CR106	1979	Tarpley JD, 1979, J Appl Meteorol, V18, P1172	59
CR107	1980	Gautier C, 1980, J Appl Meteorol, V19, P1005	74
CR108	1983	Moser W, 1983, Meteorol Rundsch, V36, P33	22
CR109	1984	Moser W, 1984, J Clim Appl Meteorol, V23, P166,	39
CR110	**1986**	**Cano D, 1986, Sol Energy, V37, P31**,	**293**
CR111	1987	Dedieu G, 1987, J Clim Appl Meteorol, V26, P79	26
CR112	1987	Raschke E, 1987, J Climatol, V7, P205,	9
CR113	1988	Diabate L, 1988, International Journal of Solar Energy, V5, P261	37
CR114	1989	Diabate L, 1989, Sol Energy, V42, P201,	34
CR115	1989	Schmetz 1989, Atmos Res, V23, P287	24
CR116	1989	Moussu G, 1989, Int J Remote Sens, V10, P1207	22
CR117	1989	Kasten F, 1989, Appl Optics, V28, P4735	14
CR118	1990	Stuhlmann R, 1990, J Appl Meteorol, V29, P586	11
CR119	1992	Pinker RT, 1992, J Appl Meteorol, V31, P194	36
CR120	1993	Noia M, 1993, Sol Energy, V51, P449	24
CR121	1993	Hay JE, 1993, Renewable Energy, V3, P381	17
CR122	1993	Noia M, 1993, Sol Energy, V51, P457	15
CR123	1996	Beyer HG, 1996, Sol Energy, V56, P207,	90
CR124	1997	Perez R, 1997, Sol Energy, V60, P89	41
CR125	1998	Ohmura A, 1998, B Am Meteorol Soc, V79, P2115	28
CR126	1998	Holben BN, 1998, Remote Sens Environ, V66, P1	21
CR127	1999	Zelenka A, 1999, Theor Appl Climatol, V62, P199	80
CR128	1999	Ineichen P, 1999, Theor Appl Climatol, V64, P119	30
CR129	1999	Hammer A, 1999, Sol Energy, V67, P139,	15
CR130	2000	Rigollier C, 2000, Sol Energy, V68, P33,	46
CR131	2002	Perez R, 2002, Sol Energy, V73, P307,	80
CR132	2003	Hammer A, 2003, Remote Sens Environ, V86, P423	85
CR133	2004	Rigollier C, 2004, Sol Energy, V77, P159	100
CR134	2004	Mueller RW, 2004, Remote Sens Environ, V91, P160	38
CR135	2005	Janjai S, 2005, Sol Energy, V78, P739	28
CR136	2009	Mueller RW, 2009, Remote Sens Environ, V113, P1012	32
CR137	2011	Blanc P, 2011, Remote Sens-Basel, V3, P343	22
CR138	2013	Sanchez-Lorenzo A, 2013, Remote Sens Environ, V134, P355,	13

The marker paper is printed in bold type

Again, the overall RPYS-CO picture can easily be divided into two time periods by the NCR values per RPY which we are going to discuss separately: the first one from the

1920 s leading to the mid-1970 s with peaks below NCR = 50 or even NCR = 20, and the second one from 1978 to 2018, with peaks increasing from NCR = 100 to NCR = 350, which are going to be in the focus of the following examination.

Time period 1922 to 1977

This period shows many peak papers we already discussed for the earlier SEM: CRs 3, 4, 6, 8, 12, 15, 16, 17, 21, 24. The most prominent peak papers are CR4 (NCR = 31) and CR6 (NCR = 18) by Angström and Prescott, respectively, as could be expected for these foundational papers. But on the other hand the SEM marker paper CR16 (Liu and Jordan 1960) is referenced only 10 times, confirming the aforementioned more specialized focus of the literature on satellite methods, making a different marker paper necessary.

After CR16 four new peak years emerge with associated papers: In 1964, Fritz et al. (CR101) from the U.S. Weather Bureau with *Satellite Measurements of Reflected Solar Energy and the Energy Received at the Ground* gives in a way the overall title for the whole satellite era of SEM. Interestingly, he *in advance* describes a method for comparing satellite with pyrheliometer measurements even before "data from accurate well-calibrated satellite instruments become available" (Fritz et al. 1964, abstract)—better than the current TIROS III data of questionable quality. In 1972, Priestley and Taylor (CR102) provided a radiation-based formula to assess heat flux and evaporation from the earth's surface, which, e.g., was much later used to evaluate satellite-derived databases as HelioClim-2 (Bois et al. 2008). In 1974, Lacis and Hansen (CR103) described a parametric method for computing the solar energy absorption on the earths's surface and in the atmosphere varying with the altitude—an important ingredient for generating solar radiation maps from satellite data (cf. CR135). In 1976, Paltridge (CR104) introduced one of the most important topics of the satellite-based methods—the short-wave albedo (diffuse reflection) of water clouds—in his case not yet measured by satellites, but by instrumented aircraft (cf. CR135).

Time period 1978 to 2018

In this focus time period of our study, we were able to identify 10 peaks at the following years: 1978-80, 1983/84, 1989/90, 1992/93, 1995/96, 1998–2000, 2002–2005, 2009, 2011, and 2013.

The only two peaks *before* the second marker paper introduce the two main methodological streams—the statistical and the physical approach—, which are able to follow through much of the total time period, assisted by the long review by Schmetz (1989) (CR115) and two shorter reviews by Noia et al. (1993a, b) (CR120, CR122).

Hay and Hanson (CR105) did a simple *empirical regression* between ground global radiance and satellite visible radiance, whereas Tarpley (CR106) from NOAA, the GOES operating agency, in addition took into account the cloudiness of the sky through a *parametrization* of the *cloud index* and *atmospheric transmittance*. The cloud index is proportional to the reflectivity of the atmosphere, and ranges from 0 at clear sky conditions to 1 at overcast (total cloud coverage). On the other side Gautier et al. (1980) (CR107) introduced their well-known "simple physical model to estimate incident solar-radiation at the surface from GOES satellite data" and evaluated it with ground measurements from Quebec.

The University of Cologne meteorologists Möser and Raschke (CR108, CR109) developed a more complex *radiative transfer model* for use with METEOSAT data. They worked out the more prominent influence of cloud coverage for solar irradiance at ground

level than of absorption by aerosols, water vapor, and ozone. The same group later applied the model with encouraging results to data of the International Satellite Cloud Climatology Project ISCCP (CR112) and was able to improve the method in 1990 (CR118). A physical model, proposed in 1987 in CR111, uses a *unique* expression for both *clear* and *cloudy conditions*, and moreover the formula of CR103 for atmospheric components.

The last peak paper on physical modeling was published in 1992 as CR119. Pinker, a meteorologist of the University of Maryland, applied her improved model from the 1980s to ISCCP data on a *global scale*.

Now we come to one of the most important statistical methods, presented in the second marker paper Cano et al. (1986) (CR110), and slightly modified by CR113, CR114, and CR116: it forms the basis for the HELIOSAT project of the Ecole des Mines de Paris, Sophia Antipolis (France), which aimed at efficiently exploiting METEOSAT images in the visible channel for solar energy applications. In so far it shares the two-fold motivation of the marker paper for SEM. The authors first deduced a *ground albedo map* from a sequence of satellite images and then estimated the *cloud index* for individual images.

Both approaches—the statistical and physical—were in a fruitful competition, which lead to more and more improvements along the way. In 1993, Hay (CR121) reviewed various modeling approaches for satellite-based estimates of solar radiation at the earth's surface, including operational techniques, which to some extent were built upon the "HELIOSAT method", before Beyer et al. (1996) (CR123) from the University of Oldenburg group were able to implement significant improvements. They accounted for some intrinsic nonlinearities of the atmosphere by using a *modeled clear sky radiation* (with absence of clouds) and gave a *geometric correction* due to the sun-satellite geometry, which is especially important for the handling of clouds in higher latitudes.

In the late 1990s, two international and wide-ranging databases for evaluation purposes had been launched: the Baseline Surface Radiation Network (BSRN), a new radiometric network of quality controlled data initiated by the World Climate Research Programme (WCRP) (CR125) and the "AERONET—A Federated Instrument Network and Data Archive for Aerosol Characterization" (CR126) initiated by NASA and using the revised optical air mass tables and approximation formula of CR117. At the same time, there were the following *evaluations* of the satellite-based methods proposed so far, based on different data from earth-based measurements: Perez from the University of Albany, NY, Zelenka from the Swiss Meteorological Institute, and co-workers undertook a comparison of GOES-8-derived radiation data with a network of 12 "ground-truth" stations in the North-eastern US concerning the site- and time-specific irradiance (CR124) and investigated the "Effective Accuracy of Satellite-Derived Hourly Irradiances" (Zelenka et al. 1999) (CR127) showing that the latter are the most accurate option beyond 25 km from a ground station. Perez and Ineichen from the University of Geneva (CR128) investigated the relationship between satellite-borne radiometer counts (from GOES and METEOSAT) and ground measurements, studying the influence of different clear sky models especially for high latitude regions with low sun elevations. Building on CR127 and CR128 they proposed in CR131 an improved and simpler version of the original model of the marker paper CR110, which is in a sense a counterpart to HELIOSAT.

From around the turn of the century, most of the peak papers have a clear bias towards European authors working on enhancing and operationalizing the HELIOSAT method:

The Oldenburg group used a statistical method to determine cloud motion vector fields for the purpose of short-term forecast of solar irradiance (CR129), which later served as one example of the application of satellite data, another being the input for the simulation of solar energy systems (CR132). The Sophia Antipolis group presented an improved

clear-sky model in CR130, which has been developed in the framework of the new digital European Solar Radiation Atlas (ESRA) that became an important ingredient of the method "HELIOSAT-2", which was no longer dependent on empirically defined parameters tuned by pyranometer measurements (CR133). Authors from the European project "HELIOSAT-3", aiming at the efficient planning and operation of solar energy systems, published a new type of solar irradiance scheme, more specific a clear-sky module called SOLIS, based on radiative transfer models using atmospheric parameter information retrieved from the current METEOSAT generation and other satellites (CR134).

The significant papers of the last three peaks in 2009, 2011, and 2013 are concerned with the benefit of satellite-based methods for *climate monitoring*: In CR136, Müller, now at the German Weather Service, and co-authors brought the physical and the statistical approaches together: the solar surface irradiance retrieval in EUMETSAT's Climate Monitoring Satellite Application Facility is based on *radiative transfer* calculations using satellite-derived parameters as input, mostly calculated using the *statistical* HELIOSAT method. In CR137, the Sophia Antipolis group reported on the HelioClim databases containing daily and monthly means of solar surface irradiance, covering Europe, Africa, and the Atlantic Ocean, that had been calculated with HELIOSAT data. Finally, CR138 gave a validation and stability assessment of radiation for the dataset of CR128 over Europe for the years 1983–2005.

Discussion

We studied the early history of SEM and the growing influence of satellite-based methods in more recent years by applying RPYS-CO to two highly cited and relevant marker papers by Liu and Jordan (1960) and Cano et al. (1986), matching the recommendations of a long-term expert. We inspected RPYs with peak citation numbers in the corresponding graphs and tables as calculated by the CRExplorer. Thereby, we were able to identify many important papers before and after the respective marker papers. They give an adequate view of most of the essential contributing streams of research in SEM, as, e.g., measurement, empirical and statistical modeling of direct and diffuse solar irradiance on the horizontal and the tilted plane on time scales from years to minutes. The topics *solar geometry, radiative transfer* calculations through the atmosphere, and *spectrally resolved* treatment of sun light can be identified as underrepresented in the results of the first RPYS-CO. But the latter of these three topics in particular gained interest only in later years and should be studied with other marker papers and the former two topics are included in the results of the second RPYS-CO—the first one in the necessary treatment of sun-satellite-geometry and the second one in the physical models.

Moreover, the second RPYS-CO has been able to trace the two methodological streams of satellite-based surface radiation retrieval from the beginning and to identify the main contributors and their competing models through the years. Finally, the operational application of the calculation schemes in near-real-time services and long-term databases, important for solar energy systems, are covered, focusing on the very recent or more distant past. The currently hot topic of forecasting solar irradiance especially for photovoltaic power production is treated only once in the peak papers (CR129), but the required time frame of several days in advance cannot be realized using satellite data only. Studying solar forecasting with RPYS-CO would only be promising with yet another marker paper.

It could be argued that an RPYS-CO on one or a few marker papers only produces a bias by favoring a limited number of research groups, maybe even enforced by the effect of self-citations. But in the given case, the first marker paper is obviously so well chosen as to unearth a diverse set of methodologies and approaches in the world-wide SEM community, coming from the two main domains *meteorology* and solar *engineering* and covering *measurement*, *modeling*, and *evaluation*. Moreover, self-citations of the repeatedly occurring (co-)authors among the peak papers play no role in our study because of their sporadic appearance.

In case of the second marker paper, stemming from a European research group heavily involved in method development, a bias on the HELIOSAT method applied to European satellites is to be observed and could be expected and may be influenced to a small extent by self-citations. But as our discussion of Table 2 makes clear, it by no means suppresses the detection of publications of competing research groups or approaches.

CR43 (Perez et al. 1987) has been suggested as an alternative SEM marker paper by the expert, but as it turned out, the RPYS-CO on both papers only confirms the results for Liu and Jordan (1960) alone (Scheidsteger and Haunschild 2019). An even stronger confirmation results from another RPYS-CO conducted for the two top most cited papers in our list of CRs together, representing the two pillars of SEM, *meteorology* and *engineering*, resp., i.e., CR34 (Collares-Pereira and Rabl 1979) with 238 citations and CR40 (Erbs et al. 1982) with 208 citations as can be seen from Fig. 6: *all* peaks and peak papers get reproduced, except from one in 1948 (CR10). This is even true for the later years after 1991 until 2013, thereby encouraging the detailed study of this period, based on these two more recent papers. But for the purpose of this paper we did not investigate the later peak structure in depth. (As a side note, the citation counts of 672 (CR34) and 515 (CR40) in the whole WoS compared to 1169 for CR16 (all numbers as of 9 April 2020) underline the good choice of the latter as marker paper.)

In total, the outcome of our study meets our expectations: All relevant historical roots of SEM research were disclosed by the RPYS-CO analysis of the first marker paper and
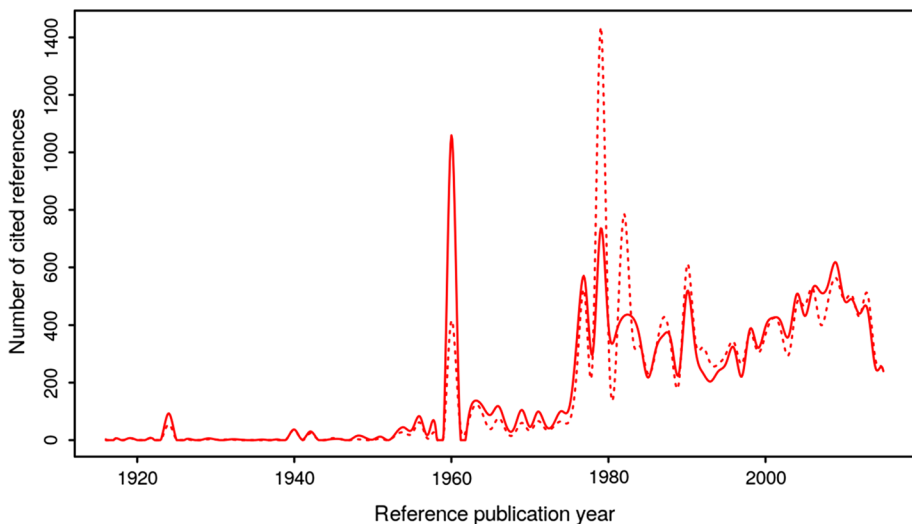
Fig. 6 Comparison of the RPYS-CO for the SEM marker paper Liu and Jordan (1960) (solid line) and the RPYS-CO for the two top cited papers (dotted line) in the corpus of papers citing Liu and Jordan (1960)

concerning the second one we are overall consistently guided through the first about 30 years of the satellite era by visiting the main contributions in terms of methodology, evaluation, and operational application. Therefore, we recommend RPYS-CO for similar investigations by researchers to get valuable insights in the development of their field of work or as a tool for studies in the history of science.

References

Arrhenius, S. (1896). On the influence of carbonic acid in the air upon the temperature of the ground. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, 41*(251), 237–276. https://doi.org/10.1080/14786449608620846.

Ballandonne, M. (2019). The historical roots (1880–1950) of recent contributions (2000–2017) to ecological economics: Insights from reference publication year spectroscopy. *Journal of Economic Methodology, 26*(4), 307–326. https://doi.org/10.1080/1350178X.2018.1554227.

Barth, A., Marx, W., Bornmann, L., & Mutz, R. (2014). On the origins and the historical roots of the Higgs boson research from a bibliometric perspective. *European Physical Journal Plus, 129*(6), 13. https://doi.org/10.1140/epjp/i2014-14111-6.

Beyer, H. G., Costanzo, C., & Heinemann, D. (1996). Modifications of the Heliosat procedure for irradiance estimates from satellite images. *Solar Energy, 56*(3), 207–212. https://doi.org/10.1016/0038-092X(95)00092-6.

Bois, B., Pieri, P., Van Leeuwen, C., Wald, L., Huard, F., Gaudillere, J.-P., & Saur, E. (2008). Using remotely sensed solar radiation data for reference evapotranspiration estimation at a daily time step. *Agricultural and Forest Meteorology, 148*(4), 619–630. https://doi.org/10.1016/j.agrformet.2007.11.005.

Bornmann, L., Haunschild, R., & Leydesdorff, L. (2018). Reference publication year spectroscopy (RPYS) of Eugene Garfield's publications. *Scientometrics, 114*(2), 439–448. https://doi.org/10.1007/s11192-017-2608-3.

Cano, D., Monget, J. M., Albuisson, M., Guillard, H., Regas, N., & Wald, L. (1986). A method for the determination of the global solar radiation from meteorological satellite data. *Solar Energy, 37*(1), 31–39. https://doi.org/10.1016/0038-092X(86)90104-0.

Chowdhury, B. H. (1990). Short-term prediction of solar irradiance using time-series analysis. *Energy Sources, 12*(2), 199–219. https://doi.org/10.1080/00908319008960199.

Collares-Pereira, M., & Rabl, A. (1979). Average distribution of solar-radiation—Correlations between diffuse and hemispherical and between daily and hourly insolation values. *Solar Energy, 22*(2), 155–164. https://doi.org/10.1016/0038-092x(79)90100-2.

Dong, B., Xu, G., Luo, X., Cai, Y., & Gao, W. (2012). A bibliometric analysis of solar power research from 1991 to 2010. *Scientometrics, 93*(3), 1101–1117. https://doi.org/10.1007/s11192-012-0730-9.

Du, H., Li, N., Brown, M. A., Peng, Y., & Shuai, Y. (2014). A bibliographic analysis of recent solar energy literatures: The expansion and evolution of a research field. *Renewable Energy, 66*, 696–706. https://doi.org/10.1016/j.renene.2014.01.018.

Dürr, B., & Zelenka, A. (2009). Deriving surface global irradiance over the Alpine region from METEOSAT Second Generation data by supplementing the HELIOSAT method. *International Journal of Remote Sensing, 30*(22), 5821–5841. https://doi.org/10.1080/01431160902744829.

Erbs, D. G., Klein, S. A., & Duffie, J. A. (1982). Estimation of the diffuse-radiation fraction for hourly, daily and monthly-average global radiation. *Solar Energy, 28*(4), 293–302. https://doi.org/10.1016/0038-092x(82)90302-4.

Fritz, S., Rao, P. K., & Weinstein, M. (1964). Satellite measurements of reflected solar energy and the energy received at the ground. *Journal of the Atmospheric Sciences*, *21*(2), 141–151.

Gautier, C., Diak, G., & Masse, S. (1980). A Simple Physical Model to Estimate Incident Solar-Radiation at the Surface from Goes Satellite Data. *Journal of Applied Meteorology*, *19*(8), 1005–1012. https://doi.org/10.1175/1520-0450(1980)019<1005:ASPMTE>2.0.CO;2.

Haunschild, R. (2019). Which are the most influential cited references in information? *Information, 10*(12), 9. https://doi.org/10.3390/info10120395.

Haunschild, R., Barth, A., & French, B. (2019a). A comprehensive analysis of the history of DFT based on the bibliometric method RPYS. *Journal of Cheminformatics*, *11*, 72. https://doi.org/10.1186/s13321-019-0395-y.

Haunschild, R., Barth, A., & Marx, W. (2016a). Evolution of DFT studies in view of a scientometric perspective. *Journal of Cheminformatics, 8*, 12. https://doi.org/10.1186/s13321-016-0166-y .

Haunschild, R., Bauer, J., & Bornmann, L. (2019b). Influential cited references in FEMS microbiology letters: Lessons from reference publication year spectroscopy (RPYS). *FEMS Microbiology Letters*, *366*(12), fnz139. https://doi.org/10.1093/femsle/fnz139.

Haunschild, R., Bornmann, L., & Marx, W. (2016b). Climate change research in view of bibliometrics. *Plos One, 11*(7), 19. https://doi.org/10.1371/journal.pone.0160393.

Haunschild, R., Leydesdorff, L., Bornmann, L., Hellsten, I., & Marx, W. (2019c). Does the public discuss other topics on climate change than researchers? A comparison of explorative networks based on author keywords and hashtags. *Journal of Informetrics, 13*(2), 695–707. https://doi.org/10.1016/j.joi.2019.03.008.

Haunschild, R., & Marx, W. (2019). *Discovering seminal works with marker papers*. Paper presented at the 8th International workshop on bibliometric-enhanced information retrieval, Cologne. http://ceur-ws.org/Vol-2345/paper3.pdf. Accessed 29 May 2019.

Haunschild, R., & Marx, W. (2020). Discovering seminal works with marker papers. *Scientometrics*. https://doi.org/10.1007/s11192-020-03358-z.

Liu, B. Y. H., & Jordan, R. C. (1960). The interrelationship and characteristic distribution of direct, diffuse and total solar radiation. *Solar Energy, 4*(3), 1–19. https://doi.org/10.1016/0038-092x(60)90062-1.

Marx, W., Bornmann, L., Barth, A., & Leydesdorff, L. (2014). Detecting the historical roots of research fields by reference publication year spectroscopy (RPYS). *Journal of the Association for Information Science and Technology, 65*(4), 751–764. https://doi.org/10.1002/asi.23089.

Marx, W., Haunschild, R., Thor, A., & Bornmann, L. (2017). Which early works are cited most frequently in climate change research literature? A bibliometric approach based on reference publication year spectroscopy. *Scientometrics, 110*(1), 335–353. https://doi.org/10.1007/s11192-016-2177-x.

McCain, K. W. (2014). Obliteration by incorporation. In B. Cronin, & C. R. Sugimoto (Eds.)*, Beyond bibliometrics: Harnessing multidimensional indicators of scholarly impact* (pp. 131–149). Cambridge, MA: MIT Press.

Michalsky, J. J. (1988). The astronomical almanac's algorithm for approximate solar position (1950–2050). *Solar Energy, 40*(3), 227–235. https://doi.org/10.1016/0038-092X(88)90045-X.

Mueller, R. W., Dagestad, K. F., Ineichen, P., Schroedter-Homscheidt, M., Cros, S., Dumortier, D., et al. (2004). Rethinking satellite-based solar irradiance modelling: The SOLIS clear-sky module. *Remote Sensing of Environment, 91*(2), 160–174. https://doi.org/10.1016/j.rse.2004.02.009.

Munkhammar, J., & Widen, J. (2016). Correlation modeling of instantaneous solar irradiance with applications to solar engineering. *Solar Energy, 133*, 14–23. https://doi.org/10.1016/j.solener.2016.03.052.

Noia, M., Ratto, C. F., & Festa, R. (1993a). Solar irradiance estimation from geostationary satellite data: I. Statistical models. *Solar Energy, 51*(6), 449–456. https://doi.org/10.1016/0038-092X(93)90130-G.

Noia, M., Ratto, C. F., & Festa, R. (1993b). Solar irradiance estimation from geostationary satellite data: II. Physical models. *Solar Energy, 51*(6), 457–465. https://doi.org/10.1016/0038-092X(93)90131-7.

Paulescu, M., Stefu, N., Calinoiu, D., Paulescu, E., Pop, N., Boata, R., & Mares, O. (2016). Angstrom–Prescott equation: Physical basis, empirical models and sensitivity analysis. *Renewable & Sustainable Energy Reviews, 62*, 495–506. https://doi.org/10.1016/j.rser.2016.04.012.

Perez, R., Ineichen, P., Seals, R., Michalsky, J., & Stewart, R. (1990). Modeling daylight availability and irradiance components from direct and global irradiance. *Solar Energy, 44*(5), 271–289. https://doi.org/10.1016/0038-092X(90)90055-H.

Perez, R., Seals, R., Ineichen, P., Stewart, R., & Menicucci, D. (1987). A new simplified version of the Perez diffuse irradiance model for tilted surfaces. *Solar Energy, 39*(3), 221–231. https://doi.org/10.1016/s0038-092x(87)80031-2.

Reindl, D. T., Beckman, W. A., & Duffie, J. A. (1990). Diffuse fraction correlations. *Solar Energy, 45*(1), 1–7. https://doi.org/10.1016/0038-092x(90)90060-p.

Scheidsteger, T., & Haunschild, R. (2019, 2–5 September 2019). *Telling the early story of solar energy meteorology by applying (co-citation) reference publication year spectroscopy.* Paper presented at the 17th International Conference on Scientometrics & Informetrics—ISSI2019—with a Special STI Indicators Conference Track.

Schmetz, J. (1989). Towards a surface radiation climatology: Retrieval of downward irradiances from satellites. *Atmospheric Research, 23*(3), 287–321. https://doi.org/10.1016/0169-8095(89)90023-9.

Stanciu, D., Stanciu, C., & Paraschiv, I. (2016). Mathematical links between optimum solar collector tilts in isotropic sky for intercepting maximum solar irradiance. *Journal of Atmospheric and Solar-Terrestrial Physics, 137*, 58–65. https://doi.org/10.1016/j.jastp.2015.11.020.

Vijayakumar, G., Kummert, M., Klein, S. A., & Beckman, W. A. (2005). Analysis of short-term solar radiation data. *Solar Energy, 79*(5), 495–504. https://doi.org/10.1016/j.solener.2004.12.005.

Yang, D., Kleissl, J., Gueymard, C. A., Pedro, H. T. C., & Coimbra, C. F. M. (2018). History and trends in solar irradiance and PV power forecasting: A preliminary assessment and review using text mining. *Solar Energy, 168*, 60–101. https://doi.org/10.1016/j.solener.2017.11.023.

Yao, Q., Li, X., Luo, F., Yang, L. P., Liu, C. J., & Sun, J. (2019). The historical roots and seminal research on health equity: A referenced publication year spectroscopy (RPYS) analysis. *International Journal for Equity in Health, 18*(1), 15. https://doi.org/10.1186/s12939-019-1058-3.

Yeung, A. W. K. (2017). Identification of seminal works that built the foundation for functional magnetic resonance imaging studies of taste and food. *Current Science, 113*(7), 1225–1227.

Yeung, A. W. K., Wong, N. S. M., & Leung, Y. Y. (2019). Are coronectomy studies being cited? A bibliometric study. *Journal of Investigative and Clinical Dentistry, 10*(1), e12366. https://doi.org/10.1111/jicd.12366.

Zelenka, A., Perez, R., Seals, R., & Renné, D. (1999). Effective accuracy of satellite-derived Hourly irradiances. *Theoretical and Applied Climatology, 62*(3), 199–207. https://doi.org/10.1007/s007040050084.