

Solar irradiation and the annual component of skin cancer incidence

Abstract

The skin cancer incidence time series is composed of principal components operating in uncorrelated time scales. This study investigates the relationship between the substantial seasonal component of skin cancer diagnoses and seasonal fluctuations in solar irradiation. After Kolmogorov-Zurbenko separation and filtration of uncorrelated time scales, cross-correlation analysis accounts for latency between irradiation and associated skin cancer detection. This study derives a coefficient of influence between changes in irradiation and skin cancer incidence, quantifying the relationship and modeling increasing risk to increased exposure. The development of this coefficient provides new opportunities to model and predict skin cancer incidence.

Keywords: skin cancer, irradiation, time series, seasonal component, kolmogorov-zurbenko filters

Volume I Issue 3 - 2014

Edward Valachovic, Igor Zurbenko

Department of Epidemiology and Biostatistics, State University of New York at Albany School of Public Health, USA

Correspondence: Edward Valachovic, Department of Epidemiology and Biostatistics, State University of New York at Albany School of Public Health, One University Place, Rensselaer, NY 12144, USA, Tel: (518) 402-0283
 Email evalachovic@albany.edu

Received: December 13, 2014 | **Published:** December 23, 2014

Abbreviations: SEER, surveillance epidemiology and end results database; OTSI, orbital total solar irradiation; GTSI, ground total solar irradiation; KZ, kolmogorov-zurbenko

Introduction

Skin cancer and total solar irradiance

Skin Cancer is the most common form of cancer.¹ The incidence of skin cancers are rising in today's populations.² Approximately 5 million will be treated for skin cancer in the United States alone.³ Prior research correlates ultraviolet exposure with increased skin cancer risk.⁴ The pronounced seasonal component in the diagnosis of skin cancer is a frequent subject of research in order to examine the effect of UV exposure.⁵ Accurately measuring a coefficient linking exposure to solar irradiation with the increase in skin cancer rates would be of immense value to research, prediction and prevention.

While possible to measure irradiation at the ground level, records on monthly or shorter time scales are susceptible to strong random regional short term fluctuations in weather patterns affecting atmospheric conditions.⁶ These changes in weather can greatly change irradiation reaching at risk populations on very short time scales. Satellite based instruments measuring total solar irradiation are not obscured by random weather patterns and their measurements may be of greater research potential in this analysis. Orbital total solar irradiation (OTSI) seasonally adjusted by latitude to reflect peak potential ground level irradiation (GTSI) is used to replicate the level of exposure at a given location averaged over many seasons.

Separation of time scales and the seasonal component

The skin cancer time series exhibits several principal component sources of variability operating at different frequencies. The seasonal component is second in total variability only to a long term increasingly upward trend. To isolate and properly investigate the seasonal component and possible relationships with the sources of variation in skin cancer incidence it is necessary to separate uncorrelated obscuring time scales such as random noise and the long term trend.⁷ Separation is achieved using a combination of low pass Kolmogorov-Zurbenko Filters.⁸ Filter parameters are selected to separate and remove the long term components and short term noise from the seasonal time scale.⁹

The principal component of variability in solar activity and irradiation occurs during at an approximate 11 year cycle, known as the solar cycle.¹⁰ Methods of time scale separation, cross-correlation, and regression previously produced a coefficient of influence for the solar cycle effect upon skin cancer prevalence.¹¹ The small but unique effect at this fingerprint frequency provided a measure of influence at this time scale upon skin cancer rates. The seasonal or one year time scale component in skin cancer by comparison is more pronounced than that associated with the solar cycle. Here seasonal patterns are not associated with fluctuations in solar activity but the axial tilt of the earth which greatly affects the levels of irradiation reaching the ground level and at risk populations.

OTSI measurements must be seasonally adjusted for irradiation reaching the ground level. This adjustment to OTSI is accomplished by compensating both for a reduction in irradiation passing through the atmosphere as well as the spread of irradiation on the earth's surface consistent with the angle of incidence given the combination of Earth's axial tilt, season and latitude.¹² Cross-correlations between skin cancer rates and irradiation must account for possible latencies both in effect and detection.¹³ Therefore cross-correlations are calculated across all possible latencies or lags in time and peak correlations are used to synchronize datasets on the candidate latency. Regression analysis is used to characterize the resulting irradiation and skin cancer incidence relationship and produces a numerical coefficient of influence. This original approach produces a coefficient appropriate for modeling and predicting skin cancer incidence at a regional level based on season, location and irradiation levels.

Methods

Data sources

Skin cancer records are extracted from monthly case diagnosis data in the SEER or Surveillance, Epidemiology, and End Results database, 1973-2010.¹⁴ The SEER sites included for this study are the states of Connecticut, Hawaii, Iowa, New Mexico, and Utah, and the cities of San Francisco-Oakland and Detroit beginning with the initiation of the database in 1973 until the most recent data published. Additional sites are available but these sites exclusively represent the earliest commencement and the longest continuous time series datasets in the SEER database. The cancer database includes all diagnoses at these

sites of each cancer type. Here, all types of skin cancer present in the database are included in the analysis. While it may be preferable to perform this analysis on individual skin cancer types with particular attention to those with greater known associations with UV exposure this study includes all types for the primary purpose of detection even if it is at the expense of eventual model fit.

Population statistics for these research sites are included in the SEER research database for analysis of rates within the observed populations. Figures are available for each observation year and monthly population changes are interpolated. Monthly skin cancer diagnosis rates were produced with these datasets for the SEER sites individually and as an aggregate group. Orbital measurements of total solar irradiation, OTSI, were recorded from the ACRIM, or Active Cavity Radiometer Irradiance Monitor, series of satellite instruments beginning in the year 1978.¹⁵ As a measure of power per unit surface area they are recorded as watts per meter squared. The ACRIM Composite is a TSI data series that is primarily composed from these instrument readings.¹⁶ Approximately 10 percent of the data of the series is missing, the longest span called the ACRIM gap between ACRIM-1 and ACRIM-2, and is filled and scaled with data from Earth Radiation Budget (ERB) experiments, Nimbus7/ERB, to relate ACRIM-1 and ACRIM-2 results. There are other composite approaches that use different subsets of satellite TSI data, ACRIM gap ratios, and different modeling, but the ACRIM Composite suits this analysis.

OTSI satellite measurements represent the solar irradiation incident upon a unit of surface area of a certain solid angle at Earth distance from the Sun. The same unit area of surface on the Earth, however, only receives a portion of the OTSI due to the atmospheric interference and the angle of incidence spreading intensity. To approximate the peak potential ground level irradiation, OTSI levels are reduced to account for these factors.¹⁷ Where θ is the latitude, there is a reduction in irradiation by a factor of cosine θ due to the spread of intensity over a wider surface of the Earth at the given latitude. This is compounded by a further cosine θ reduction due to the thickness of the atmosphere through which the irradiation traverses at the given latitude. Thus, a fraction of OTSI equal to the cosine squared of the site's latitude represents the maximum reaching the ground level.

Analysis methods

Analysis is performed in R version 3.1.1 statistical software using the KZA package.¹⁸ The raw data is transformed by the natural logarithm function to stabilize variance of observations across time (Figure 1). The time series is a compilation of the signal energies from different components that operate in distinct time scales and are therefore uncorrelated. Analysis in a particular time scale requires the separation and filtering of those that are uncorrelated but may interfere. Kolmogorov-Zurbenko filters are low pass filters characterized by two parameters. With notation KZ (km) they are k iterations of a standard moving average filter of m points defining the moving average filter window.¹⁹ With interest in the seasonal time scales, they are well suited in this analysis for separation and filtering.¹² Here, it is necessary to remove both long term (trend) and short term (noise) time scales. A KZ (12,3) filter is used to separate the longer term (>1 year) components. In this case, these longer term components are removed from the original signal, making a high pass filter, retaining signals of 12 months or less. It is to this band width that a KZ (3,3) filter is then applied thereby filtering short time scale

noise. Filtering produces and edge effect in the data therefore these observations are discarded based on the size of the filter window and number of iterations.

Cross-correlation or correlating between data with a lag helps identify latency in possible causative relationships. At the peak in cross-correlations the data sets are paired to compensate for any latency caused either in effect or detection. Accounting for this candidate latency, the lagged seasonal component of skin cancer incidence is then regressed over the seasonal component of ground level TSI producing a slope coefficient, or coefficient of influence.

Results

Time scale separation

The natural logarithm transformed datasets exhibit the characteristics of their influential components. The skin cancer rate long term component exhibits an upward trend characteristic with growth over time (Figure 2). The seasonal component corresponding to the time scale surrounding 1 year exhibits a cyclic pattern of peaks and troughs with the highest skin cancer rates occurring during summer months and diminished rates during winter months consistent with what might be anticipated for the northern latitudes. The very short term (<4 months) exhibits the random fluctuations in monthly observations characteristic of noise.

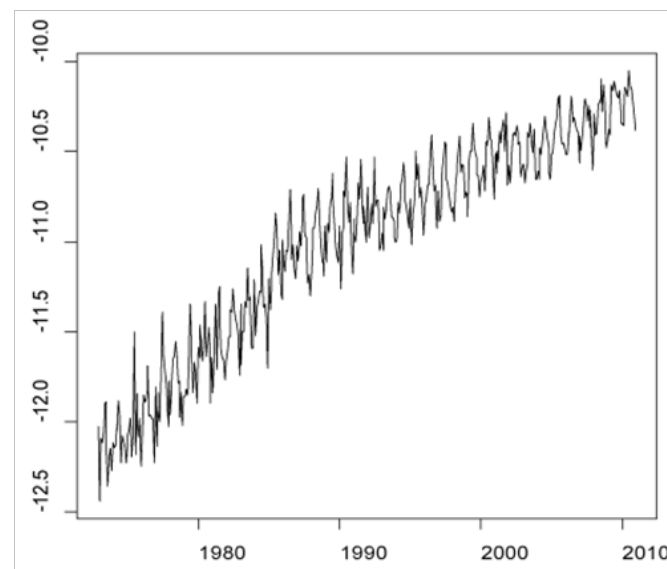


Figure 1 Log Skin Cancer Monthly Rate.

The principal component in OTSI corresponds with the natural fluctuation of solar intensity characteristic of the solar cycle. The Solar Cycle is the approximate 11 year cycle of changing solar activity that affects many solar characteristics including but not limited to, total solar irradiation, magnetic field strength, and sunspot activity.²⁰ The influence of the 11 year solar cycle upon long term skin cancer rates operates at a different frequency and is uncorrelated with the seasonal time scale. The natural logarithm transformed OTSI dataset exhibits the characteristic solar cycle in the long term time scale, and negligible variation around the seasonal time scale (Figure 3). This is anticipated given that solar activity is under insignificant influence from Earthly seasonal cycles. Short term time scales again exhibit the random variations of noise.

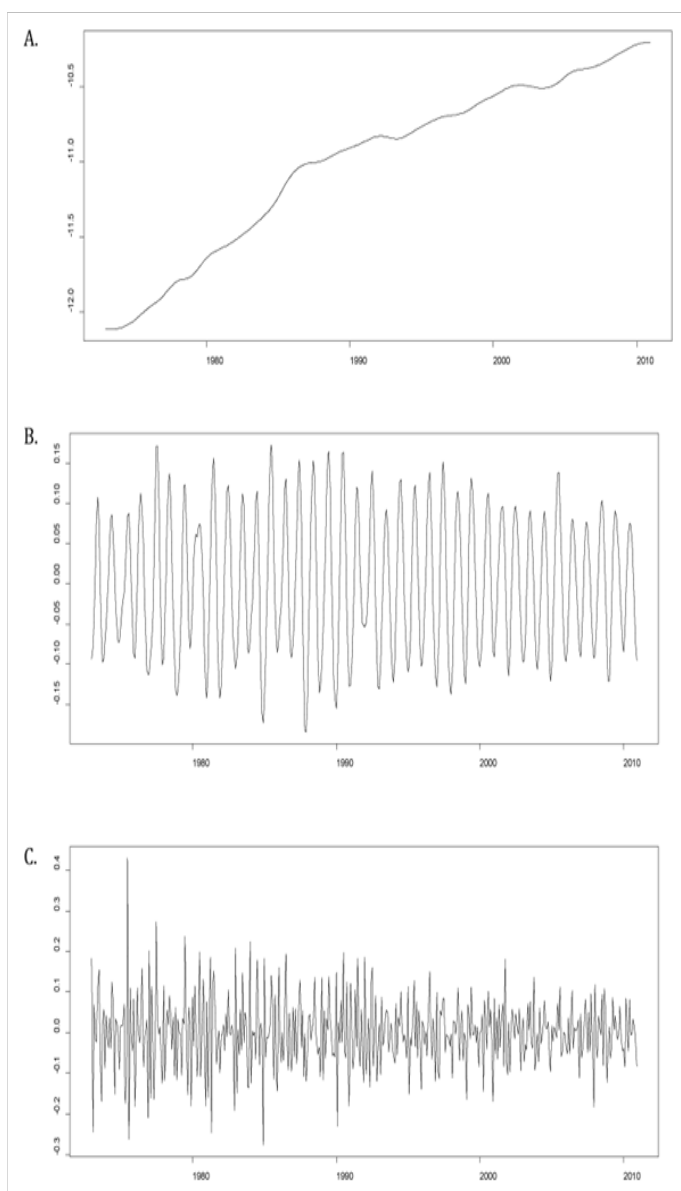


Figure 2 Log Skin CancerRate (a) Long Term Trend (b) Seasonal Component, and (c) Noise.

After adjusting for latitude, ground total solar irradiance, GTSI, unlike OTSI does display a strong seasonal component of peaks and troughs consistent with the seasonal changes in solar intensity at the surface level elevations for the northern hemisphere. This seasonal ground effect caused by the earth's axial tilt dominates all other factors of variation in GTSI by many magnitudes, making both the OTSI variations from the solar cycle variation and noise nearly invisible given the scale (Figure 4).

Cross-correlations

Monthly observations of the seasonal component of GTSI show the anticipated peak around June corresponding to the most direct surface exposure in the northern hemisphere toward the Sun (Figure 5).

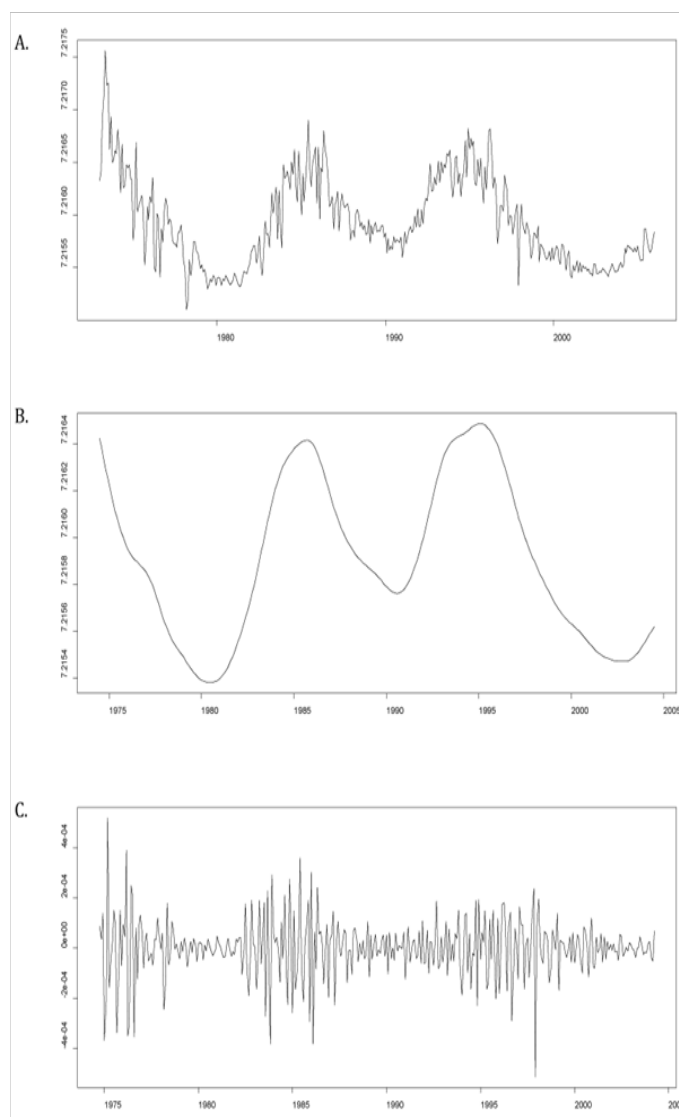


Figure 3 Log OTSI (a) Raw Data (b) Long Term (> 1 year), (c) Noise.

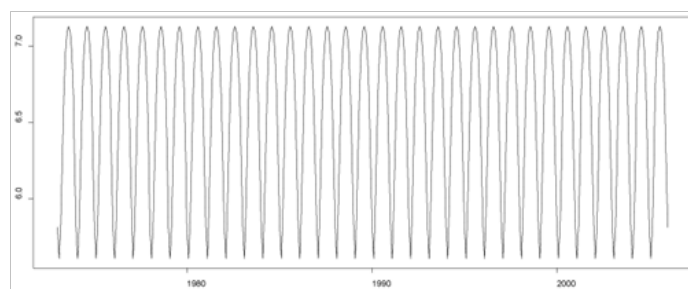


Figure 4 Middle US latitude adjusted Log GTSI.

Monthly observations of the annual component of skin cancer rate peak in July collectively among the SEER sites, a lag of one month behind the peak in GTSI (Figure 6). Individually, SEER sites are split between those with peaks in June or July, a lag of 0 or 1 month, indicating a true lag may be of the order of a fraction of a month. However, such precision is finer than the scale of measure in the datasets.

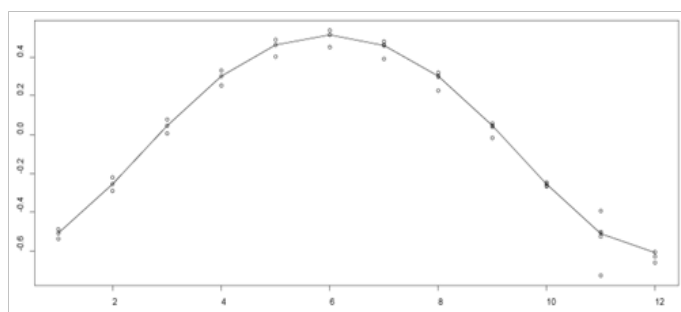


Figure 5 Monthly observations and mean curve of the seasonal component of Log GTSI.

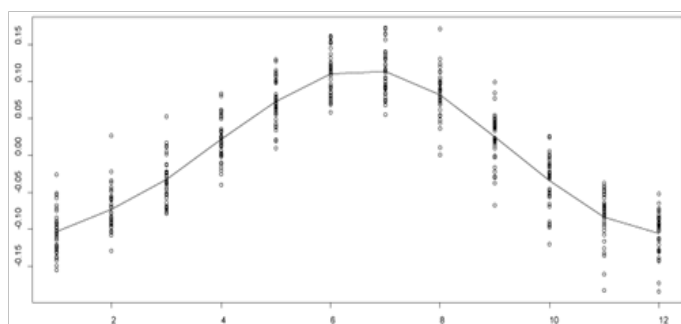


Figure 6 Monthly observations and mean curve of the seasonal component of Log Skin Cancer Rate.

When using the 1 month latency between seasonal GTSI and skin cancer incidence the cross-correlation maximizes at a value of .920. Regressing seasonal skin cancer rate on seasonal GTSI with this 1 month latency results in an estimate of the slope coefficient, or coefficient of influence, of approximately 0.202 and an R-Squared, the square of correlation, of 0.85 (Figure 7).

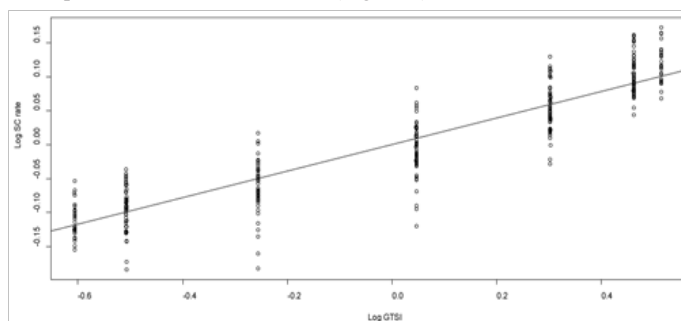


Figure 7 Ln Seasonal Skin Cancer (SC) Rate vs. Ln Seasonal GTSI with a 1 month latency.

The aggregated and the individual SEER site results include results from regression at both 0 and 1 month of lag between GTSI adjusted for each individual site's latitude and individual site skin cancer rate (Table 1).

These SEER site coefficients of influence that result from regressions at 1 month latency are plotted against site latitude (Figure 8). Regression on these observations produces a slope coefficient of latitude -0.0055, and an R-Squared 0.6107. With data limited to these few sites there does not appear any strong evidence of a pattern based upon geography or latitude, with the exception of Hawaii. Although this outlying observation is insufficient to draw conclusions it raises

questions of geographic influence, such as the cosine squared of latitude multiplier used in this study.^{12,17}

$$b_i : N(0, D)$$

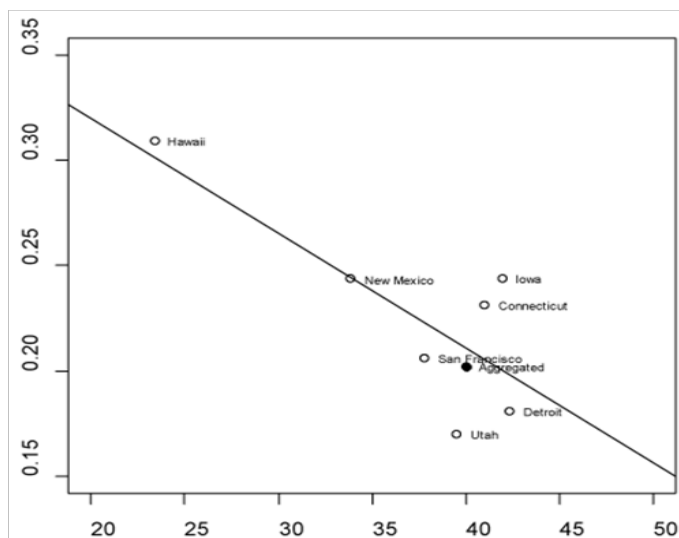


Figure 8 Coefficients of Influence at 1 Month Latency vs. Latitude.

Discussion

While the existence of a relationship between skin cancer and solar irradiation exposure is clear, the unique strength of this analytical method is the derivation of a numerical coefficient of influence between the two. The time series decomposition approach as compared to cancer risk studies makes explanation more complicated, but it provides high coefficients of explanation that inevitably will give invaluable clues to disease modeling and preventive measures. This is true for the seasonal components of this analysis as well as the overall skin cancer rate, which is one possible extension of this analysis. This also provides the prospect of extending the analysis into forecasting and prediction.

The coefficient of influence between seasonal components of ground adjusted total solar irradiation (GTSI) and skin cancer rate is approximately 0.20. Noting the Log scale, this translates into every 10 percent increase in seasonal GTSI being associated with an increase in seasonal skin cancer incidence of approximately 2 percent. However, an average season at middle US latitudes can include summer irradiance that is three times as strong as winter irradiance. This corresponds to an approximate 25 percent increase in the rate of skin cancer diagnosis from the winter low to the summer high. Furthermore, results produced an R-Squared of .8461, which as a measure of fit indicates that seasonally changing GTSI explains approximately 85 percent of the seasonal variability in skin cancer incidence accounting for one month of latency.

Skin cancer is associated with an accumulation of exposure over years or even decades²¹. Here, the seasonal effect and associated 1 month or less latency does not address this long term latency. The seasonal effect may be indicative of a triggering mechanism just prior to detection. The mechanism and cause may be observational in nature such as surpassing some diagnostic threshold or some other means that causes detection. This is a subject for continued research.

Individual site results that cross-correlate and regress against

their respective regionally latitude adjusted level of GTSI produce a range of coefficients from a low of 0.17 in Utah to a high of 0.31 in Hawaii. Interpreting these coefficients, they correspond to a range from 1.6 to 3 percent increase in skin cancer incidence for a 10 percent

increase in GTSI. The coefficients may relate to a total background level at each site created by long term spatial radiation and midterm climate fluctuations.^{12,17} Multivariate analysis of the corresponding components with different variables may reveal the truth.

Table 1 Site coefficients of influence

Site	Coefficient (0 months lag)	R-Squared	Coefficient (1 month lag)	R-Squared	Middle Latitude
Aggregated	0.1991	0.8191	0.2022012	0.8461	40
Connecticut	0.218808	0.6352	0.231368	0.7102	41.017
San Francisco	0.1959000	0.6465	0.2062795	0.718	37.783
Detroit	0.1969120	0.6715	0.1808	0.5723	42.333
Hawaii	0.3152264	0.2954	0.3095760	0.2866	23.367
Iowa	0.2322942	0.6439	0.2441210	0.7115	41.942
New Mexico	0.2642	0.4465	0.2436928	0.3800	33.833
Utah	0.166515	0.4621	0.1699463	0.4849	39.5

This analysis can be extended in several other ways. Foremost would include additional research sites. More sites are available through the SEER research database although for shorter time spans. This posed a true obstacle when investigating the long term trends and solar cycle component of skin cancer incidence, but would be less of a hindrance when investigating the seasonal component.¹¹ The inclusion of additional sites raises the prospect of examining the relationship between latitude and site specific coefficients of influence. Inclusion of additional sites around the globe would also extend these results outside of these US geographic ranges to include different northern latitudes and test the result for a reversed seasonal effect on southern latitudes.

This analysis can extend to specific skin cancer types thereby focusing on those with greater known solar exposure risk factors. Also the inclusion of more refined data, such as site specific factors like weather that modifies exposure or population specific factors like race and gender would strengthen this analysis by adjusting for confounding factors. This may lead to more precise disease specific coefficients of influence, modeling and prediction possibilities.

Acknowledgement

None.

Conflict of interests

Authors declare that there is no conflict of interest.

References

1. Stern RS. Prevalence of a history of skin cancer in 2007: results of an incidence-based model. *Arch Dermatol*. 2010;146(3): 279–282.
2. Karia PS, Han J, Schmultz CD. Cutaneous squamous cell carcinoma: estimated incidence of disease, nodal metastasis, and deaths from disease in the United States, 2012. *J Am Acad Dermatol*. 2013;68(6): 957–966.
3. Guy GP, Machlin SR, Ekwueme DU, et al. Prevalence and costs of skin cancer treatment in the US, 2002–2006 and 2007–2011. *Am J Prev Med*. 2014;S0749–3797(14): 00510–00518.
4. Koh HK, Geller AC, Miller DR, et al. Prevention and early detection strategies for melanoma and skin cancer: Current status. *Arch Dermatol*. 1996;132(4):436–443.
5. Keller AK, Uter W, Pfahlberg AB, et al. Seasonality of cutaneous melanoma diagnoses: a comprehensive comparison of results in Bavaria and Northern Ireland. *Melanoma Res*. 2013;23(4):321–330.
6. Eskridge R, Ku J, Rao S, et al. Separating different scales of motion in time series of meteorological variables. *Bull Amer Meteor Soc*. 1997;78:1473–1483.
7. Tsakiri KG, Zurbenko I. Effect of noise in principal component analysis. *Journal of Statistics and Mathematics*. 2011;2(2):40–48.
8. Kolmogorov–Zurbenko Filter. In Wikipedia: The Free Encyclopedia. Wikimedia Foundation Inc. Encyclopedia on–line.
9. Yang W, Zurbenko I (2010) Kolmogorov–Zurbenko Filters. *Computational Statistics*. 2014;2(3):340–351.
10. Wang WM, Lean J, Sheeley NJ. Modeling the Sun's Magnetic Field and Irradiance since 1703. *The Astrophysical Journal*. 2005;625:522–538.
11. Valachovic E, Zurbenko I. Skin Cancer, Irradiation, and Sunspots: The Solar Cycle Effect. *Biomed Res Int*. 2014;538574.
12. Zurbenko I, Cyr D. Climate fluctuations in time and space. *Climate Research*. 2011;57(1):93–94.
13. Nadler D, Zurbenko I. Estimating Cancer Latency Times Using a Weibull Model. *Advances in Epidemiology*. 2014;8.
14. Surveillance, Epidemiology, and End Results (SEER) Program (www.seer.cancer.gov) Research Data (1973–2010), National Cancer Institute, DCCPS, Surveillance Research Program, Surveillance Systems Branch,

- released April 2013, based on the November 2012 submission.
15. <http://www.acrim.com/>
 16. Willson RC, Mordvinov AV. Secular total solar irradiance trend during solar cycles 21–23. *Geophysical Research Letters*. 2003;30(5): 1199.
 17. Zurbenko I, Ming L. Restoration of Time–Spatial Scales in Global Temperature Data. *AJCC*. 2012;1(3):154–163.
 18. Close B, Zurbenko I. Kolmogorov–Zurbenko adaptive filters. (Version 3), 2013.
 19. Zurbenko I. The Spectral Analysis of Time Series. North–Holland, 1986;pp. 248.
 20. Hempelmann A, Weber W. Correlation Between the Sunspot Number, the Total Solar Irradiance, and the Terrestrial Insolation. *Solar Physics*. 2012;277(2): 417–430.