**SCIBASE Journals**

**SciBase Oncology**

# Cardio-Cancer Risks Cross-Validation by Gaidai-Yakimov Reliability Method

Oleg Gaidai[1]; Vladimir Yakimov[2]; Kelin Wang[1*]

[1]Shanghai Ocean University, Shanghai, China.

[2]Central Marine Research and Design Institute, Saint Petersburg, Russia.

**Corresponding Author: Kelin Wang**

Shanghai Ocean University, Shanghai, China.

Email: m210811423@st.shou.edu.cn

**Citation:** Gaidai O, Yakimov V, Wang K. Cardio-Cancer Risks Cross-Validation by Gaidai-Yakimov Reliability Method. SciBase Oncol. 2023; 1(2): 1007.

## Abstract

This study validates novel biological and health systems reliability method, in particular suitable for complex multi-dimensional public health and biological systems, versus the well-established bivariate bio-reliability method, being known to accurately assess bivariate dynamic system's extreme contours. Contemporary reliability methods, that deal with spatio-temporal observations, do not always have advantage of coping easily with high-dimensional biosystems, along with complex cross-correlations various different system components. The suggested methodology coped with this challenge well. This study aimed to assess joint cardiovascular along with cancer diseases death rate risks, at any given time horizon, and in any region/country of interest, by means of applying novel Gaidai-Yakimov reliability method to raw clinical dataset. Multicentre, population-based, spatio-temporal, medical survey dataset analysis, based on novel bio-statistical method has been adopted. In this study, annual numbers of deaths, associated with cancer and cardiovascular diseases, recorded in all 195 world's countries, have been selected. It has always been challenging to model such phenomena, due to bio-system non-stationarity, high-dimensionality, and cross-couplings between different biological or public health system's components. Main purpose of this study was to benchmark Gaidai-Yakimov state-of-the-art technique, enabling optimal and efficient extraction of relevant statistical information from an underlying raw clinical dataset. Suggested methodology may be used for prognostics in various public digital health applications, based on their available clinical raw survey datasets.

## Introduction

CVD (Cardiovascular disease) is a term being used to refer to a range of human diseases, affecting blood vessels and heart. Those include abnormal blood pressures (hypertensions); Heart attacks (coronary heart diseases); Strokes (cerebrovascular diseases); Heart failures; other numerous related heart disorders. NCI (National Cancer Institute) defines cancers as collection of serious diseases, when abnormal cells start-actively divide, and then spread to neighbouring tissues. Cancers may arise within various parts of human body resulting in a wide range of cancer types, in some cases spreading to other key parts of human body, through lymph and blood systems. CVDs along with cancers being 2 most major mortality reasons around the globe.

Considering public system's associated global risks, it makes sense to study bivariate statistics of both CVDs as well as cancers. This study considers public health and bio system, but it does not address health issues of individual patients. This study intends to contribute to contemporary public health and bio system research, but not clinical one; potential rationale could be an assessment of relative burdens on public health within various countries, given time-horizon of interest.

Statistical aspects of CVD and cancer diseases risks, along with the statistically relevant methods, has been receiving extensive attention in modern research community [1-14]; for CVD-related issues see [15-22]; for cancer-related issues see

[33-51]. Generally, it being quite challenging to assess risk factors and reliability of realistic bio and health systems, as well as to predict excessive in situ mortality risks, under actual CVD and cancer disease conditions, using contemporary theoretical reliability and statistical methods [1,23-26]. Latter being typically due to bio and health system's high-dimensionality, large number of random factors, governing bio and health system's dynamics, yet system's spread over extensive terrains. In general, reliability of complex biological, environmental or health systems may be assessed straightforwardly, by having either enough measurements, or by performing extensive direct MC (Monte Carlo) simulations. For CVD along with cancer diseases, available observational clinical datasets are often quite limited both in time, locations, as well as in dataset sizes, [1]. Given above-described challenges, authors have introduced here novel Gaida-Xing reliability methodology for bio and health systems, able to predict accurately CVD and cancer diseases excessive mortality risks. This study has been focused on CVD and cancer diseases spread dynamics over the globe [1,4,22], with special focus on complex cross correlations between various countries. Whole world has been hosen because of extensive public health observations, as well as widely available related research.

Reliability and statistics provide basis for lifetime modelling, typically using EVT (Extreme Value Theory), as a popular approach in both medical research and bio-medical engineering. In [15], authors proposed novel approach, utilizing PVF (Power Variance Function) copulas (Clayton, inverse Gaussian, Gumbel copulas, etc.), along with conditional sampling, used for survival analysis. For recent studies, discussing upper bounds distribution of life expectancy [16,19]. As there being not much statistical research done to assess contemporary risks associated with CVD and cancer diseases, including numerous heart and blood vessels diseases, authors introduced in this study novel reliability method, able to analyse excessive mortality episodes and their spatiotemporal spread, providing better indications and insight of possible future above-mentioned diseases spread. In our study CVD and cancer disease excessive mortality episodes have been viewed as unexpected/random incidents, that may possibly occur in any region/country, at any time-horizon, hence spatial spread being well accounted for. Non-dimensional scaling factor λ has been introduced, in order to predict latter CVD and cancer disease risks, given temporal return period and specific spatial area of interest.

Bio and health systems being subjected to various ergodic and cyclic environmental influences. Bio-process can be viewed as being intricately 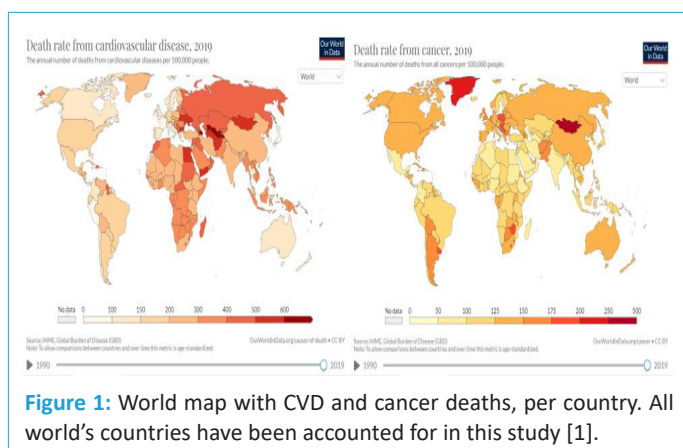dependent on a number of bio/environmental parameters, whose temporal variation may be approximated/modelled, as ergodic or cyclic processes on their own.



**Figure 2:** Left: Cardio-cancer's relationship illustration [2] Right: biomarkers [3].

The incidence dataset of CVD and cancer diseases within one hundred-ninety-five world countries, during three recent decades 1990-2019 has been od frobtainm a public website [1]. This clinical dataset being given per each world's country, the bio or health system, under consideration may be regarded as MDOF (Multi Degree of Freedom) dynamic bio-system, having highly cross-correlated national/regional components (or dimensions). This study assists managing global risks of future CVD and cancer diseases excessive mortality episodes, by prognosticating them, and hence it is being only focused on annually globally registered patient death-numbers, not on disease symptoms. All world countries have been accounted for in this study [1] (Figure 1).

Although studied dataset being only limited cardio-cancer dataset, the proposed bio-reliability technique may well be applied for any well documented diseases.

Heart failure and cancer being the 2 main causes of death in developed countries. Even though these 2 clinical entities seem to be distinct, they really share many of the same risk factors, symptoms, and pathophysiological pathways (inflammation, activation of the immunological and neurohormonal systems, metabolic abnormalities, endothelial dysfunction, etc.). Along with the well-known cardiotoxic effects of onco and cardiac treatment, cancers and heart failure are assumed to be related by a bi-directional interaction, with one disease favoring the other and vice versa (Figure 2).

### Gaidai-Yakimov method

Consider MDOF public health or biological system's vector (X(t),Y(t),Z(t),…), composed from bio or public health system critical/key components X(t),Y(t),Z(t),… that has been either observed, measured or numerically simulated over a sufficiently long (say, representative) observation time-period (0,T). Unidimensional (1D) bio or health system component's global maxima denoted here as,

$$X_T^{\max} = \max_{0 \le t \le T} X(t)$$

$$Y_T^{\max} = \max_{0 \le t \le T} Y(t)$$

$$Z_T^{\max} = \max_{0 \le t \le T} Z(t),$$

with sufficiently long (representative) time lapse T we basically mean large T, with respect to the bio-dynamic system auto-correlation, and relaxation times. Let $X_1,…,X_{N_X}$ be consequent in time local maxima of the biosystem component process X=X(t) at discrete, temporally increasing times



**Figure 1:** World map with CVD and cancer deaths, per country. All world's countries have been accounted for in this study [1].
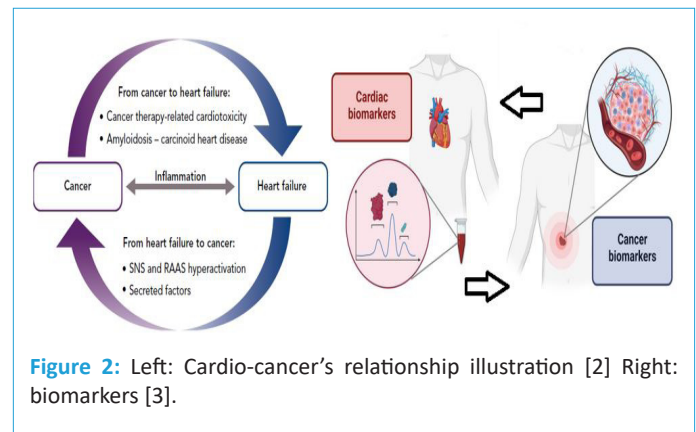
$$t_1^X < \cdots < t_{N_X}^X$$

within (0,T). Identical type of definitions to follow for other MDOF components Y(t),Z(t),… Namely $Y_1,…,Y_{N_Y}$); $Z_1,…,Z_{N_Z}$ etc. For the sake of simplicity, all bio or health system's local maxima have been assumed to be positive. Then

$$P = \iiint_{(0,\ 0,\ 0,\ ,…)}^{(\eta_X,\ \eta_Y,\ \eta_Z,\ …)} p_{X_T^{\max},\ Y_T^{\max},\ Z_T^{\max},…}\left(x_T^{\max}, y_T^{\max}, z_T^{\max}, …\right) dx_T^{\max} dy_T^{\max} dz_T^{\max}\ … \quad (1)$$

being probability of dynamic biosystem's survival, having critical (hazard) values of biosystem components, denoted as $\eta_X, \eta_Y, \eta_Z,…$; U being the logical unity operator «or»;

$$p_{X_T^{\max},\ Y_T^{\max},\ Z_T^{\max},…}$$

being biosystem's joint PDF (probability density function) of individual biosystem component's maxima. When biosystem's NDOF (Number of Degrees of Freedom) being large, it being not always practically feasible to directly assess biosystem's joint PDF

$$p_{X_T^{\max},\ Y_T^{\max},\ Z_T^{\max},…}$$

and hence biosystem's probability P of survival. Latter probability P needs to be accurately assessed, as bio or health system's expected lifetime, following Eq. (1). Bio-system's 1D components X,Y,Z,… being re-scaled, as well as non-dimensionalized in a following manner

$$X \to \frac{X}{\eta_X},\ \ Y \to \frac{Y}{\eta_Y},\ \ Z \to \frac{X}{\eta_X},\ \ … \quad (2)$$

making all bio-system's components non-dimensional, having now the same failure/hazard limits $\lambda=1$. Next, 1D biosystem's components local maxima being merged into 1D temporally non-decreasing synthetic biosystem's vector

$$R(t) \equiv \vec{R} = (R_1, R_2, …, R_N)$$

following corresponding merged time-vector $t_1 \leq … \leq t_N$, $N \leq N_X + N_Y + N_Z + …$ Every local maxima $R_j$ being actually encountered by one of bio-system components: either X(t) or Y(t), or
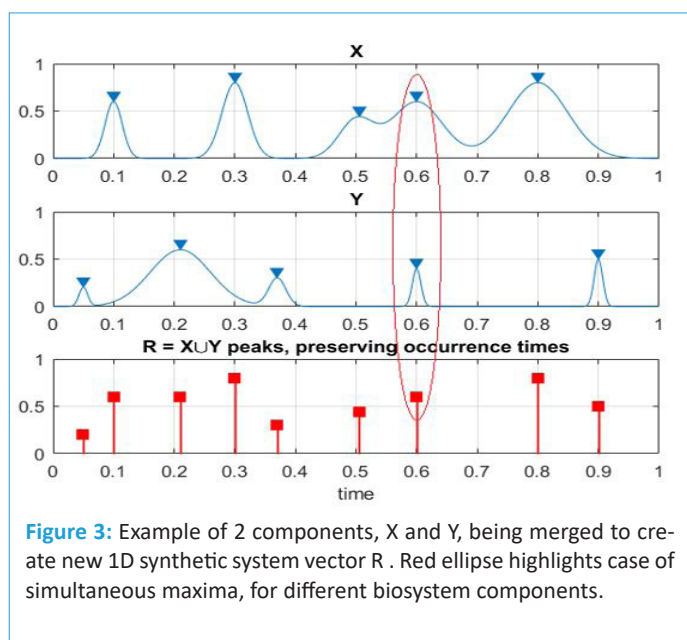


**Figure 3:** Example of 2 components, X and Y, being merged to create new 1D synthetic system vector R . Red ellipse highlights case of simultaneous maxima, for different biosystem components.

Z(t). Synthetic biosystem $\vec{R}$ -vector has 0 data-loss, see Figure 3.

Now the non-decreasing synthetic biosystem vector $\vec{R}$, along with it's corresponding non-decreasing occurrence times

$t_1 \leq … \leq t_N$, have been thus introduced, [13].

## Results

This section applies the Gaidai-Yakimov approach, [61] to a bivariate (2D) Random bio-process Z(t)=(X(t),Y(t)). This bivariate process being made up of annually registered patient death numbers, X(t),Y(t), for CVD and cancer-related diseases respectively, which have been monitored synchronously, within representative time-lapse (0,T). Assume that N equidistant discrete temporal intervals $t_1,…,t_N$ within observational time-lapse were used to collect samples (data points) $(X_1,Y_1),…,(X_N,Y_N)$ within (0,T). In this study, we investigate bivariate joint CDF (Cumulative Distribution Function) $P(\xi,\eta):= \text{Prob}\ (X^{˘}_N \leq \xi, Y^{˘}_N \leq \eta)$ of the 2D vector $(X^{˘}_N, Y^{˘}_N)$, having components $X^{˘}_N = \max\{X_j\ ; j=1,…,N\}$, and $Y^{˘}_N = \max\{Y_j\ ; j=1,…,N\}$. The latter example serves as an illustration of a dynamic two-dimensional (2D) biosystem. Critical (hazard) Thresholds were determined, using unidimensional extreme bio or health system component's global maxima with specific return periods and risks/probabilities p. Following scaling has been done to combine both clinical/measured/observed time-series X,Y following Eq. (2), Resulting in each of 2 biosystem components having the same failure/hazard limits equal to 1. Hence, by maintaining assembled local maxima in temporally increasing order, all bio or health system component's local maxima, from each observed/measured time-series, were combined into a single 1D system's vector $\vec{R}$=(max$\{X_1,Y_1\}$,…,max$\{X_N,Y_N\}$).
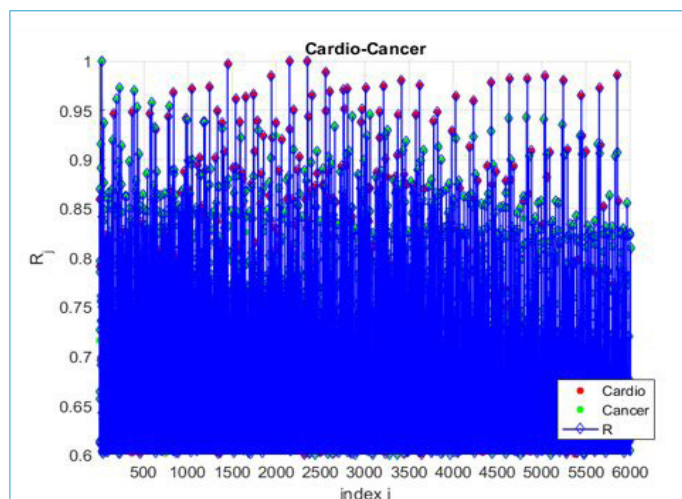


**Figure 4:** Example of cardio-cancer data, visualized as time-series.
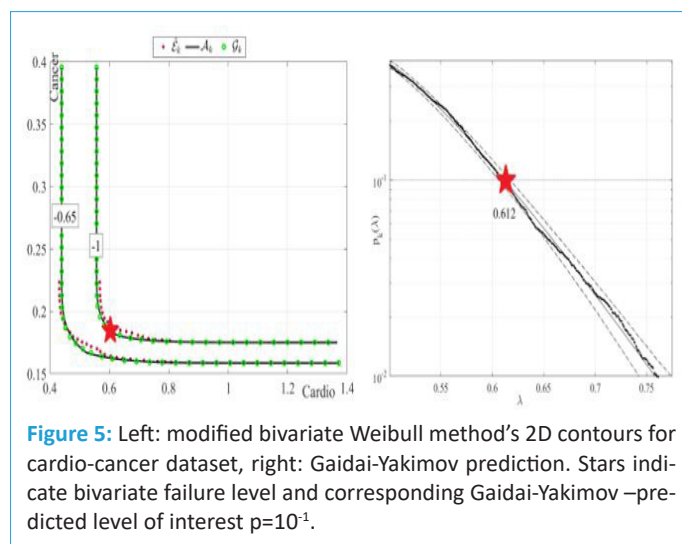


**Figure 5:** Left: modified bivariate Weibull method's 2D contours for cardio-cancer dataset, right: Gaidai-Yakimov prediction. Stars indicate bivariate failure level and corresponding Gaidai-Yakimov –predicted level of interest $p=10^{-1}$.

Figure 4 presents an example of cardio-cancer data, given as time-series.

Figure 5 presents modified bivariate Weibull method's bivariate (2D) contours for cardio-cancer dataset. It being seen from Figure 5 there being an intrinsic inaccuracy, owing to the specific copula choice within the modified bivariate Weibull method, as it fits several Gumbel copulas to the measured clinical dataset. For more information on the modified bivariate Weibull methods, see [5-9]. Selected test-point being located on the $p=10^{-1.5}$ contour line, and it was accurately verified by the modified bivariate Weibull method, given the bivariate failure point $\eta\_X=44000$, $\eta\_Y=65$. Then, confirming estimate from Gaidai-Yakimov approach has been compared to the same probability level $p=10^{-1.5}$, matching the latter 2D contour line. It was seen that Gaidai-Yakimov method's predicted 95% CI (Confidence Interval) for the modified bivariate Weibull method probability level estimate was indeed accurate, [10-14]. Note that the above-described methodology, while unique, has practical benefits of effectively utilizing measured/observed/simulated datasets, since it can handle biosystem's multi-dimensionality, and conduct accurate extrapolation, based on relatively scarce raw clinical datasets.

Poincare plot being defined as SODP (Second Order Difference Plot). Time-series based dataset's quality may be observed/checked statistically, using SODP, by analysing underlying dataset's successive differences. SODP assists statistical pattern recongnition through consecutive differences in the underlying raw time-series data. CEM (Circled Entropy Measurement) method calculated data fall over circled regions, see Figure 6. MATLAB script for feature SODP models has been shared for practitioners [62].
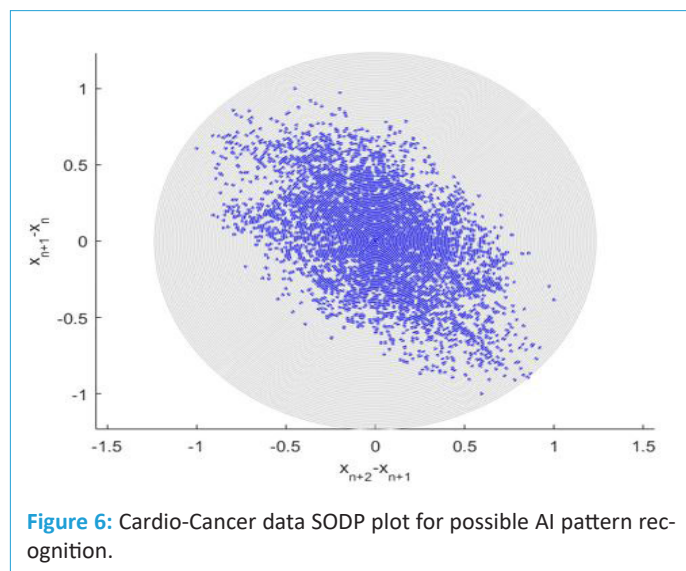


**Figure 6:** Cardio-Cancer data SODP plot for possible AI pattern recognition.

Figure 6 presents 2nd order SODP plot, such type of plots may be used to identify raw clinical dataset patterns, and compare them with other datasets, for example within a framework of entropy-based artificial intelligence (AI) recognition methods [52]. Figure 6 highlights relationship between annual registered CVD and cancer patient recorded death numbers. As an example, SODP pattern, extracted from reduced dataset can be compared with SODP pattern, extracted from full dataset, and certain judgement about reduced da-taset representativity and data quality may be drawn. While current study presents the MDOF biosystem's reliability study by sub-asymptotic method, the commonly used EVT being asymptotic and being only 1DOF.

Classic contemporary health systems reliability approaches do not have advantage of easily dealing with observed clinical time-series, originating from complex bio-systems with high-dimensionality and non-linear cross correlations between various bio-system's components. The key advantages of Gaidai-Yakimov method, proposed here, being its capacity to analyze reliability and risks of high-dimensional non-linear dynamic bio and public health systems [53-60]. Despite apparent simplicity, the current work presents essentially novel multi-dimensional bio-modelling methodology, along with methodological route to apply epidemic/chronical disease forecasting, while bio or health system has not yet reached its critical/hazard levels.

### Conclusions

The advantages of working with highly dimensional systems and complicated cross-correlation between different bio-system components readily are not provided by traditional dependability methodologies for clinical time series. The primary advantage of the suggested technique is its ability to analyze dependability of high-dimensional dynamic systems. In this study, time series representing the rates of cardiovascular and cancer death from accessible clinical datasets were presented. The theoretical support for the proposed approach has been developed. The complexity and high dimensionality of dynamic bio and health systems require the development of novel, accurate, yet robust techniques that can handle the limited clinical datasets that are currently available and utilize clinical data as effectively as possible, despite the appeal of using direct measurements or extensive Monte Carlo simulations to analyze bio-system's reliability. Only for one-dimensional dynamic systems, the method proposed in this article has previously been shown useful when used with a range of simulation models. In general, forecasts turned out to be rather accurate. This study has focused on a multi-dimensional bio-reliability strategy that is general-purpose, trust worthy, and user-friendly. The suggested method can be applied in a number of biological, medicinal, and health sectors. The offered clinical example in no way limits the potential uses of the suggested technique.

### Declarations

**Ethical approval**: Not applicable.

**Conflict of interests**: Authors declare that they have no conflict of interest.

**Authors' contributions**: All authors contributed equally.

**Data availability**: Datasets used and analysed during the current study available from the corresponding author on reasonable request.

### References

1. Our world in data, https://ourworldindata.org/causes-of-death.

2. Bertero E, Ameri P, Maack C. Bidirectional Relationship Between Cancer and Heart Failure: Old and New Issues in Cardio-oncology, Card Fail Rev. 2019; 5: 106-11. doi: 10.15420/cfr.2019.1.2.

3. Chianca M, Panichella G, Fabiani I, et al. Bidirectional Relationship Between Cancer and Heart Failure: Insights on Circulating Biomarkers, Frontiers in Cardiovascular Medicine. 2022; 2022: 9. https://doi.org/10.3389/fcvm.2022.936654.

4. Gaidai O, Cao Y, Xing Y, Wang J. Piezoelectric Energy Harvester Response Statistics, Micromachines. 2023; 14(2): 271. https://doi.org/10.3390/mi14020271.

5. Gaidai O, Wang F, Wu Y, Xing Y, Medina A, Wang J. Offshore renewable energy site correlated wind-wave statistics, Probabilistic Engineering Mechanics. 2022; 2022(68). https://doi.org/10.1016/j.probengmech.2022.103207.

6. Gaidai O, Yan P, Xing Y. Future world cancer death rate prediction, Scientific Reports. 2023; 13(1): 10.1038/s41598-023-27547-x.

7. Gaidai O, Xu J, Hu Q, Xing Y, Zhang F. Offshore tethered platform springing response statistics, Scientific Reports, Vol. 12, www.nature.com/articles/s41598-022-25806-xGaidai, O, Xing Y, Xu, X, 2022, COVID-19 epidemic forecast in USA East coast by novel reliability approach, Research square. 2022. DOI: https://doi.org/10.21203/rs.3.rs-1573862/v1

8. Xu X, Xing Y, Gaidai O, Wang K, Sandipkumar Patel K, Dou P, Zhang Z. A novel multi-dimensional reliability approach for floating wind turbines under power production conditions, Frontiers in Marine Science. 2022. https://doi.org/10.3389/fmars.2022.970081.

9. Gaidai O, Xing Y, Balakrishna R. Improving extreme response prediction of a subsea shuttle tanker hovering in ocean current using an alternative highly correlated response signal, Results in Engineering. 2022. https://doi.org/10.1016/j.rineng.2022.100593.

10. Cheng Y, Gaidai O, Yurchenko D, Xu X, Gao S. Study on the Dynamics of a Payload Influence in the Polar Ship, The 32nd International Ocean and Polar Engineering Conference, Paper Number: ISOPE-I-22-342. 2022.

11. Gaidai, O, Xing, Y. A Novel Multi Regional Reliability Method for COVID-19 Death Forecast, Engineered Science. 2022. DOI: 10.30919/es8d799.

12. Gaidai O, Xing Y. A novel bio-system reliability approach for multi-state COVID-19 epidemic forecast, Engineered Science, DOI: 10.30919/es8d797Gaidai, O, Fu, S, Xing Y. Novel reliability method for multidimensional nonlinear dynamic systems, Marine Structures. 2022; 2022(86). https://doi.org/10.1016/j.marstruc.2022.103278.

13. Gaidai O, Yan P, Xing Y. Future world cancer death rate prediction, Scientific Reports. 2023; 13(1): 10.1038/s41598-023-27547-x.

14. Gaidai O, Xu J, Yan P, Xing Y, Zhang F, Wu Y. Novel methods for wind speeds prediction across multiple locations, Scientific reports. 2022; 12: 19614. https://doi.org/10.1038/s41598-022-24061-4.

15. Tsao C, et al. Heart Disease and Stroke Statistics-2022 Update: A Report from the American Heart Association, Circulation, Vol. 2022; 145(8). https://doi.org/10.1161/CIR.0000000000001052.

16. Smolina K, Wright FL, Rayner M, Goldacre MJ. Determinants of the decline in mortality from acute myocardial infarction in England between 2002 and 2010: linked national database study. BMJ 2012; 344: 8059.

17. Balakrishna, R, Bjørnerud, T, Bemanian, M, Aune, D, & Fadnes, L. T. Consumption of Nuts and Seeds and Health Outcomes Including Cardiovascular Disease, Diabetes and Metabolic Disease, Cancer, and Mortality: An Umbrella Review. Advances in Nutrition. 2022.

18. Mackay DF, Irfan MO, Haw S, Pell JP. Meta-analysis of the effect of comprehensive smoke-free legislation on acute coronary events. 2010; 96: 1525-1530.

19. Cox B, Vangronsveld J, Nawrot TS. Impact of stepwise introduction of smokefree legislation on population rates of acute myocardial infarction deaths in Flanders, Belgium. 2014; 100: 1430-1435.

20. Alzuhairi KS, Søgaard P, Ravkilde J, Gislason G, Køber L, Torp-Pedersen C. Incidence and outcome of first myocardial infarction according to gender and age in Denmark over a 35-year period (1978-2012). Eur Heart J Qual Care Clin Outcomes. 2015; 1: 72-78.

21. Mirzaei M, Truswell AS, Taylor R, Leeder SR. Coronary heart disease epidemics: not all the same. 2009; 95: 740-746.

22. NCDRF Collaboration. Trends in adult body-mass index in 200 countries from 1975 to 2014: a pooled analysis of 1698 population-based measurement studies with 19.2 million participants. Lancet 2016; 387: 1377-1396.

23. Madsen HO, Krenk S, Lind NC. Methods of structural safety. Englewood Cliffs, Prentice-Hall Inc. 1986.

24. Ditlevsen O, Madsen HO. Structural reliability methods. Chichester (World): John Wiley & Sons, Inc. 1996.

25. Melchers RE. Structural reliability analysis and prediction. New York, John Wiley & Sons, Inc. 1999.

26. Choi S-K, Grandhi RV, Canfield RA. Reliability-based structural design, London, Springer-Verlag. 2007.

27. Pickands J. Statistical inference using extreme order statistics. Ann Stat. 1975; 3(1): 119-131.

28. Zhao X. Extreme value modelling with application in finance and neonatal research. PhD Thesis, The University of Canterbury. Available. 2010. http://ir.canterbury.ac.nz/bitstream/10092/4024/1/thesis_fulltext.pdf.

29. Zheng L, Ismail K, Meng XH. Freeway safety estimation using extreme value theory approaches: A comparative study. Accident Anal Prev. 2014; 62: 32-41. doi: 10.1016/j.aap.2013.09.006 PMID:24129319.

30. McNeil AJ, Frey R, Embrechts P. Quantitative risk management. Concepts, techniques and tools.America. Princeton University Press. 2005.

31. Patie P. Estimation of value at risk using extreme value theory. 2000 In: Talks in financial and insurance mathematics [Internet]. LaWorldnne Eidgenossische Technische Hochschule Zürich. 1855. [about 1 screen]. Available: http://www.math.ethz.ch/*patie/VaREvT.pdf, 2000.

32. Songchitruksa P, Tarko Andrew P. The extreme value theory approach to safety estimation. Accident Anal Prev. 2006; 38: 811-822. PMID: 16546103.

33. Siegel R, Miller K, Fuchs H, Jemal A. Cancer statistics, CA: A Cancer Journal for Clinicians. 2022. https://doi.org/10.3322/caac.21708.

34. Yabroff KR, Wu, X C, Negoita S, Stevens J, Coyle L, Zhao J, et al. Association of the COVID-19 Pandemic with Patterns of Statewide Cancer Services. J Natl Cancer Inst. 2021.

35. Surveillance, Epidemiology, and End Results (SEER) Program. SEER*Stat Database: Incidence- SEER 9 Registries Research Data with Delay- Adjustment, Malignant Only, November 2020 Submission (1975- 2018) <Katrina/Rita Population Adjustment>- Linked to County Attributes- Total U.S, 1969- 2018 Counties. National Cancer Institute, Division of Cancer Control and Population Sciences, Surveillance Research Program, Surveillance Systems Branch; 2021.

36. Surveillance, Epidemiology, and End Results (SEER) Program. SEER*Stat Database: Incidence- SEER 18 Registries Research Data + Hurricane Katrina Impacted Louisiana Cases, November 2020

Submission (2000- 2018) <Katrina/Rita Population Adjustment>- Linked to County Attributes- Total U.S, 1969-2018 Counties. National Cancer Institute, Division of Cancer Control and Population Sciences, Surveillance Research Program, Surveillance Systems Branch. 2021.

37. Surveillance Research Program. SEER*Explorer: an interactive website for SEER cancer statistics. National Cancer Institute; 2021. Accessed. 2021. seer.cancer.gov/explorer/.

38. Surveillance, Epidemiology, and End Results (SEER) Program. SEER*Stat Database: Incidence- SEER Researc Limited-Field Data With Delay- Adjustment, 21 Registries, Malignant Only, November 2020 Submission (2000- 2018)- Linked To County Attributes- Time Dependent (1990- 2018) Income/Rurality, 1969-2019 Counties. National Cancer Institute, Division of Cancer Control and Population Sciences, Surveillance Research Program. 2021.

39. Surveillance Research Program, Statistic Methodology and Applications. DevCan: Probability of Developing or Dying of Cancer Software. Version 6.7.9. National Cancer Institute. 2021.

40. Surveillance, Epidemiology, and End Results (SEER) Program. SEER*Stat Database: North American Association of Central Cancer Registries (NAACCR) Incidence Data- Cancer in North America Analytic File, 1995-2018, With Race/Ethnicity, Custom File With County, American Cancer Society Facts and Figures Projection Project (which includes data from the Center for Disease Control and Prevention's National Program of Cancer Registries, the Canadian Council of Cancer Registries' Provincial and Territorial Registries, and the National Cancer Institute's SEER Registries, certified by the NAACCR as meeting high-quality incidence data standards for the specified time periods). National Cancer Institute, Division of Cancer Control and Population Sciences, Surveillance Research Program; 2021. 8. Sherman R, Firth R, Charlton M, et al, eds. Cancer in North America: 2014-2018. Volume One: Combined Cancer Incidence for the United States, Canada and North America. North American Association of Central Cancer Registries, Inc. 2021.

41. Sherman R, Firth R, Charlton M, et al, eds. Cancer in North America: 2014-2018. Volume Two: Registry- Specific Cancer Incidence in the United States and Canada. North American Association of Central Cancer Registries, Inc. 2021.

42. Surveillance, Epidemiology, and End Results (SEER) Program. SEER*Stat Database: Mortality- All Causes of Death, Total U.S. (1969- 2019) <Katrina/Rita Population Adjustment>- Linked To County Attributes - Total U.S, 1969- 2019 Counties (underlying mortality data provided by the National Center for Health Statistics). National Cancer Institute, Division of Cancer Control and Population Sciences, Surveillance Research Program; 2021.

43. Wingo PA, Cardinez CJ, Landis SH, et al. Long- term trends in cancer mortality in the United States, 1930- 1998. Cancer. 2003; 97(12): 3133-3275.

44. Murphy SL, Kochanek KD, Xu J, Heron M. Deaths: Final Data for 2012. National Vital Statistics Reports. Vol 63, No. 9. National Center for Health Statistics. 2015.

45. Steliarova- Foucher E, Stiller C, Lacour B, Kaatsch P. International Classification of Childhood Cancer, third edition. Cancer. 2005; 103: 1457-1467.

46. Fritz A, Percy C, Jack A, et al, eds. International Classification of Diseases for Oncology. 3rd ed. World Health Organization. 2000.

47. World Health Organization (WHO). International Statistical Classification of Diseases and Related Health Problems, 10th revision. Vol I- III. WHO. 2011.

48. Surveillance Research Program. SEER*Stat software, version 8.3.8. National Cancer Institute. 2020.

49. Surveillance Research Program. Joinpoint Regression Program version 4.9.0.1. National Cancer Institute. Statistical Research and Applications Branch. 2021.

50. Mariotto AB, Zou Z, Johnson CJ, Scoppa S, Weir HK, Huang B. Geographical, racial and socio- economic variation in life expectancy in the US and their impact on cancer relative survival. PLoS One. 2018; 13: 0201034.

51. Clegg LX, Fever EJ, Mistune DN, Fay MP, Hankey BF. Impact of reporting delay and reporting error on cancer incidence rates and trends. J Natl Cancer Inst. 2002; 94: 1537-1545.

52. Yayık A, Kutlu Y, Altan G. Regularized HessELM and Inclined Entropy Measurement for Congestive Heart Failure Prediction. 2019. Cornell University, https://arxiv.org/abs/1907.05888.

53. Gaidai O, Cao Y, Loginov S. Global cardiovascular diseases death rate prediction, Current Problems in Cardiology. 2023. https://doi.org/10.1016/j.cpcardiol.2023.101622.

54. Gaidai O, Cao Y. Xing Y. Balakrishna R. Extreme springing response statistics of a tethered platform by deconvolution, International Journal of Naval Architecture and Ocean Engineering. 2023. https://doi.org/10.1016/j.ijnaoe.2023.100515.

55. Gaidai O, Xing Y, Balakrishna R, Xu J. Improving extreme offshore wind speed prediction by using deconvolution, Heliyon. 2023. https://doi.org/10.1016/j.heliyon.2023.e13533.

56. Gaidai O, Xing Y. Prediction of death rates for cardiovascular diseases and cancers, Cancer Innovation. 2023. http://doi.org/10.1002/cai2.47.

57. Gaidai O, Cao Y, Xing Y, Balakrishna R. Extreme springing response statistics of a tethered platform by deconvolution, International Journal of Naval Architecture and Ocean Engineering. 2023. https://doi.org/10.1016/j.ijnaoe.2023.100515.

58. Gaidai O, Xing Y, Balakrishna R, Xu J. Improving extreme offshore wind speed prediction by using deconvolution. 2023. Heliyon, https://doi.org/10.1016/j.heliyon.2023.e13533.

59. Gaidai O, Xing Y. Prediction of death rates for cardiovascular diseases and cancers, Cancer. 2023. Innovation, http://doi.org/10.1002/cai2.47.

60. Gaidai O, Wang F, Yakimov V. COVID-19 multi-state epidemic forecast in India. Proceedings of the Indian National Science Academy. 2023. 10.1007/s43538-022-00147-5.

61. Gaidai O, Xing Y, Xu J, Balakrishna R. Gaidai-Yakimov reliability method validation for 10-MW floating wind turbines. Scientific Reports. 2023; 13(1). DOI: 10.1038/s41598-023-33699-7.

62. https://github.com/apdullahyayik/time-series-analysis