Information of Joint and Marginal Probability Distributions

Kalman Ziha
Faculty of Mechanical Engineering and Naval Architecture, Ivana Lucica 5, 10000 Zagreb
kziha@fsb.hr

Abstract. *The paper in the first place reviews the information and the uncertainty measures of joint and marginal probability distributions of the sets and subsets of random events. Next it reminds on the relations of the unconditional and conditional entropy of joint and marginal distributions and their combinations. Then it elaborates the ways how these measures can be applied in the sea surface uncertainty evaluation, that is, how to gain the information from the wave properties visually observed and brought together in the Global Wave Statistics. Furthermore, the examples offered demonstrate the computational procedures for the sea surface uncertainty evaluation and prospective usefulness of the information obtainable from the joint and marginal distributions of observed explicit and combined wave properties.*

Keywords. information, uncertainty, entropy, probability, waves, Global Wave Statistics

1. Introduction

The paper exposes how the entropy concept, which emerged earlier in the information theory [1][2][3] for the evaluation of the amount of information and was later generalized in the probability theory [4][5][6] for an objective uncertainty assessment of systems and subsystems of events, can be applied to the joint and marginal probability distributions. The event oriented system analysis that combines the probability and information theory for engineering purposes [7][8][9][10] tackles the sea surface uncertainty by using the joint probability distributions of observable wave properties collected in the Global Wave Statistics (GWS) [11]. Although the entropy of a system of events is an objective property, since it depends only on the events themselves [5], the nature of human perception involved in visual observations introduces subjectivity into the evaluation of the sea surface uncertainty. In despite of the subjectivity of visual wave observations, the paper supports the belief in an appropriate ordering of uncertainties founded on the entropy and on the coherent data in the GWS.

2. The distribution of wave properties

The sets of visually observable N_h=15 significant wave heights and N_t=11 zero crossing wave periods in any of the N_d=8 or all wave directional classes *d*, Fig. 1, are the basic sources of uncertainty that embrace information of the sea surface alienated in N_M=104 areas *A* for N_s=4 seasons in the GWS [11], Fig. 2.

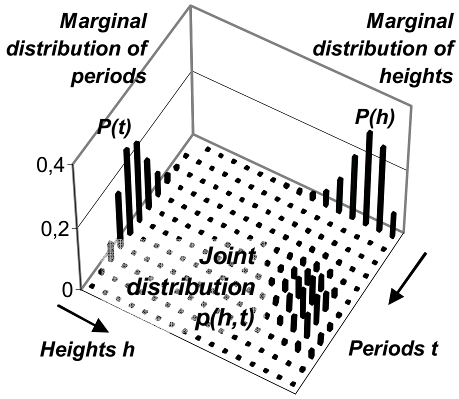

Figure 1. Wave scatter diagram of joint and marginal distribution of heights and periods for a directional class o fan area in GWS

The annual and four seasonal subdivisions are denoted *s=(annual), March-May, June-August, September-November, December-February* and the wave directional classes are denoted *d=(all), NW, N, NE, W, E, SW, S, SE* or combinations, for example *A25:MM,NW*. Of particular interest might be the joint observations in an area *A* collected together on annual basis and in all directions *d*. The joint probability distribution $p(h,t)$ is known as the wave scatter diagram, Fig. 1. It is the entry to tabular calculations of different aspects of information and uncertainties og GWS in parts per thousands of observations of directional class *d* of an area *A* and season *s*.

The condition of completeness of $N_{ht}=N_h \times N_t$ observable properties observed properties of a directional class *d* in an area *A* is given by:

$$\sum_{all\ h}\sum_{all\ t} p(h,t) = 1 \qquad (1)$$

The logarithms in next examples are of base two and the entropy is expressed in *bits*.

Proceedings of the *ITI 2007 29th Int. Conf. on Information Technology Interfaces,* June 25-28, 2007, Cavtat, Croatia

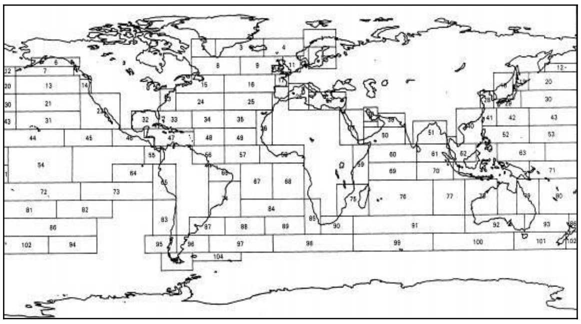

Figure 2. Definitions of areas in GWS [11]

3. Unconditional entropy

The information or the uncertainty of a directional class d can be assessed by the unconditional Shannon's entropy [3] taking the reported joint probability distribution of all N_{ht} significant wave heights and zero crossing periods from the table of directional class d [11] as a finite scheme of events [4] of a complete set of observations [5], e.g. Table 1, as shown:

$$H(d) = -\sum_{all\ h}\sum_{all\ t} p(h,t)\cdot \log p(h,t) \qquad (2)$$

The entropy (2) in *bits* and the average number of events (*epe*) $F(d) = 2^{H(d)}$ indicate the information and uncertainty of a directional class d with respect to both heights and periods. The maximally entropy of a directional class is:

$$H_{\max}(d) = \log N_h \cdot N_t = 7{,}366\, bits \text{ or } F(d) = 165\, epe$$

Besides the uncertainty of all joint observations in a directional class of interest are the separate uncertainties of marginal distributions of particular significant wave heights h_w and zero crossing periods t_z, respectively, Fig. 1. The marginal probability distributions for height h_w and period t_z are:

$$P(h_w) = \sum_{all\ t} p(h_w, t) \qquad (3)$$

$$P(t_z) = \sum_{all\ h} p(h, t_z) \qquad (4)$$

The unconditional entropy (2) of a directional class, e.g. Table 1, with respect to the marginal distributions of all heights and periods, respectively, are defined using (3) and (4), as:

$$H(d^h) = -\sum_{all\ h} P(h)\cdot \log P(h) \qquad (5)$$

$$H(d^t) = -\sum_{all\ t} P(t)\cdot \log P(t) \qquad (6)$$

The entropy (5, 6) and the average numbers $F(d^h) = 2^{H(d^h)}$ and $F(d^t) = 2^{H(d^t)}$ indicate the information and the uncertainty of the marginal distribution of heights and periods.

4. Conditional entropy

Instead of the unconditional entropy of the marginal distribution of heights and periods as it was elaborated earlier in section 2, a more comprehensive insight to uncertainties but with more computational effort can be obtained.

The conditional entropy [5][7] of the directional class d with respect to a selected wave height h_w as well as relative to periods t_z, Table 1, viewed as subsystems of events, is defined as follows:

$$H(d/h_w) = -\sum_{all\ t} \frac{p(h_w,t)}{P(h_w)} \log \frac{p(h_w,t)}{P(h_w)} \qquad (7)$$

Instead of the unconditional entropy $H(d^h)$ of marginal distribution of heights (5) the conditional entropy of the directional class d relative to all wave heights h is as follows:

$$H(d/h) = \sum_{all\ h} P(h_w)\cdot H(d/h_w) \qquad (8)$$

The following relations are valid for periods:

$$H(d/t_z) = -\sum_{all\ h} \frac{p(h,t_z)}{P(t_z)} \log \frac{p(h,t_z)}{P(t_z)} \qquad (9)$$

$$H(d/t) = \sum_{all\ t} P(t_z)\cdot H(d/t_z) \qquad (10)$$

The relations (7-9) follow directly from the theorem of the mixtures of distributions [5].

5. The relation of entropy

The following relation among the joint distributions of a directional class d and marginal distributions of heights h and periods t can be proven starting of the definition of the unconditional Shannon's entropy of wave directional classes d (2) relative to marginal probability distributions of significant wave heights $h=h_w$, using the notation in text (5, 8), (analogous is for the marginal distributions of zero crossing periods using (6, 10)), as shown:

$$\begin{aligned} H(d) &= \\ &= -\sum_{all\ h}\sum_{all\ t} [P(h_w) p(h_w,t)]\cdot \log[P(h_w) p(h_w,t)] = \\ &= \sum_{all\ h} P(h_w)\cdot H(d/h_w) - \sum_{all\ h} P(h_w)\log P(h_w) = \\ &= H(d/h) + H(d^h) \end{aligned} \qquad (11)$$

6. Combinations of properties

The entropy of a combination c of heights $h(c)$ of a directional class d in an area $A{:}s$ with overall probability $p_{d^{h(c)}} = \sum_{all\ h(c)} P(h)$, is defined as shown:

$$H(d^{h(c)}) = \frac{1}{p_{d^{h(c)}}} \sum_{h(c)} H(d^h) + \log p_{d^{h(c)}} \qquad (12)$$

The combination (12) uses the partial results $H(d_{h(c)})$ of $H(d^h)$ (5).

And analogously is the combination of periods $t(c)$ of a directional class d:

$$H(d^{t(c)}) = \frac{1}{p_{d^{t(c)}}} \sum_{t(c)} H(d^t) + \log p_{d^{t(c)}} \qquad (13)$$

Starting from the definition of entropy of a subsystem [8] composed of the combination $h(c)$ of marginal distribution of heights (5) in a directional class d, it follows (analogous is for zero crossing periods) the proof:

$$H(d^{h(c)}) = -\sum_{h(c)} \sum_{all\,t} \frac{P(h_w)}{p_{d^{h(c)}}} \log \frac{P(h_w)}{p_{d^{h(c)}}} =$$

$$= -\frac{1}{p_{d^{h(c)}}} \sum_{h(c)} P(h_w) \cdot \log P(h_w) +$$

$$+ \frac{\log p_{d^{h(c)}}}{p_{d^{h(c)}}} \sum_{h(c)} P(h_w) =$$

$$= \frac{1}{p_{d^{h(c)}}} \sum_{h(c)} H(d^h) + \log p_{d^{h(c)}}$$

7. Examples

The unconditional entropy (2) of $N_{ht}=73$ joint observations of the area *A9:annual,all* in North Atlantic, Fig. 2 and Table 1, is as follows:

$$H(A9{:}s = anual, d = all) =$$

$$= -\sum_{all\ h} \sum_{all\ t} p_{A9,d}(h,t) \cdot \log p_{A9,d}(h,t) = 5{,}087\ bits$$

$h=H/\log N_{ht}=0{,}854$, $F(A25{:}ann,all)= 33{,}98$ *epe*.

The entropy of N_h=11 observed heights (5) and of N_t=10 observed periods (6) of the marginal probability distributions, Table 1, are as shown:

$H(d^h) = 2{,}786\ bits$ or $F(d^h)=6{,}896\ epe$

$H(d^t) = 2{,}570\ bits$ or $F(d^t)=5{,}940\ epe$.

Particularly, the conditional entropy of heights (8) and periods (10), Table 1, are as follows:

$H(d/h) = 2{,}301\ bits$ or $F(d/h)=4{,}927\ epe$

$H(d/t) = 2{,}517\ bits$ or $F(d/t)=5{,}7212\ epe$.

The following example evaluates the combined uncertainty with respect to selection c of heights $h(c)>3 meters$ of significant wave heights for area $A9{:}annual,d=all^{h>3}$. The probability to encounter waves $h>3$ *meters* amounts to $p_{A29,d=all^{h>3}} = 0{,}498$. Employing the partial results in Table 1, from (12) it follows:

$$H(d^{h>3}) = \frac{1{,}602}{0{,}498} - 1{,}006 = 2{,}211\ bits$$

or

$F(d^{h>3})$=4,430 *epe*.

Table 1. Unconditional and conditional entropy of joint and marginal distributions of wave heights (meters) and periods (seconds) in Area 9 (North Atlantic) all directions on annual basis

	$H(d/t_z)$		0,863	1,755	2,037	2,209	2,372	2,517	2,578	2,703	2,156	0,000	2,517	$H(d/t)$
	$P(t_z)$	Σ=1	0,002	0,024	0,114	0,241	0,274	0,196	0,098	0,037	0,012	0,002	2,570	$H(d^t)$
	>14								t_z					
	13-14						$H(d)$=5,088 bits						$P(h_w)$	$H(d/h_w)$
	12-13												Σ=1	
1	11-12						$p(h,t)$		0,001				0,001	0
3	10-11							0,001	0,001	0,001			0,003	1,585
5	9-10						0,001	0,002	0,002	0,001	0,001		0,007	2,236
6	8-9					0,001	0,002	0,003	0,003	0,002	0,001		0,012	2,459
6	7-8					0,001	0,004	0,006	0,005	0,003	0,001		0,02	2,328
8	6-7				0,001	0,003	0,009	0,012	0,009	0,004	0,002	0,001	0,041	2,557
8	5-6				0,001	0,009	0,02	0,022	0,015	0,006	0,002	0,001	0,076	2,443
7	4-5				0,004	0,021	0,04	0,037	0,021	0,008	0,002		0,133	2,363
8	3-4			0,001	0,012	0,045	0,066	0,05	0,022	0,007	0,002		0,205	2,357
8	2-3			0,003	0,028	0,075	0,08	0,045	0,015	0,004	0,001		0,251	2,29
7	1-2			0,009	0,048	0,073	0,048	0,017	0,004	0,001			0,2	2,174
6	0-1		0,002	0,011	0,02	0,013	0,004	0,001					0,051	2,092
N_t	$h(m)/t(s)$	<4	4-5	5-6	6-7	7-8	8-9	9-10	10-11	11-12	12-13	>13	2,786	2,301
73	N_h	0	1	4	7	9	10	11	11	10	8	2	$H(d^h)$	$H(d/h)$

The example indicates that the sailing in North Atlantic (Area 9) with respect to wave heights and periods is comparable to the uncertainty of 33,98 equally probable events or *epe*. This uncertainty is appropriate to gambling with slightly less of 34 and notably more than with 33 cards. Such a presentation of uncertainty by 33,98 *epe* is hopefully more perceptible for seamen than by 5,087 bits.

8. Conclusions

The paper employs the entropy concept as defined in the probability theory to investigate the uncertainties of joint and marginal distributions. It appears that to each joint probability distribution not only the entropy of marginal distributions can be assigned but also the entropy of the joint distribution conditional on the marginal distributions. Moreover, the entropy of the joint distribution, the entropy of the marginal distributions and the conditional entropy of joint distribution relative to marginal distributions are mutually related.

The examples demonstrate how the visual observations reported in the GWS can be used not only for probability estimates but also for the assessment of the information and uncertainty of the observable sea surface properties. The entropy concept in the probability theory can appropriately indicate the uncertainty of joint and marginal distributions of reported significant wave heights and zero crossing periods. Consequently, the sea surface can be characterized not only by the physical properties such as the wave energy distribution but also by the uncertainty that corresponds to the information contained in the reported probability distributions of the visually observable wave properties. The human experience of uncertain world is commonly related to hazardous games and different kinds of gambling. The amount of information of one *bit* reflects the uncertainty of two equally probable events that one faces when tossing a coin for example. The paper holds that for practical purposes is the representation of uncertainty by numbers of equally probable events intuitively more appropriate. Whether the entropy is an objective property of the sea surface or a subjective entity that requires conscious observers, this is still an unresolved issue - mostly a problem for philosophers.

Nevertheless, the paper supports the entropy as an appropriate mathematical concept that can define the amount of information or the uncertainties embraced in the sets of visual wave observations. The coherent ordering of uncertainty in the paper relies on the GWS unified data organization that can provide insight to the world wide sea surface uncertainty distribution.

9. References

[1] [1] Hartley, R V, Transmission of Information, Bell Systems Tech.J, 7, 1928.

[2] Wiener, N, Cybernetic, or Control and Comm., Bell System Tech. J, 27, 1948.

[3] Shannon, C E, Weaver, W, The Mathematical Theory of Communication, Urbana Univ. Illinois, 1949.

[4] Khinchin, A I, Mathematical Foundations of Information Theory, Dover Publications, NY, 1957.

[5] Renyi, A, Probability Theory, North-Holland, Amsterdam, 1970.

[6] Aczel, J, Daroczy, Z, On Measures of Information and Their Characterization, Acad. Press, NY, 1975.

[7] Ziha, K, Event Oriented System Analysis, Probabilistic Eng. Mech., 15(3), 2000.

[8] Ziha, K, Entropy of Subsystems of Events, In: Kalpic D, Hljuz Dobric V, editors. Proceedings of the 20th International Conference on Information Technology Interfaces; Pula, Croatia. Zagreb: SRCE University Computing Centre, University of Zagreb; 1998.

[9] Ziha, K, Usage of Relative Uncertainty Measures, In: Kalpic D, Hljuz Dobric V, editors. Proceedings of the 21th International Conference on Information Technology Interfaces; Pula, Croatia. Zagreb: SRCE University Computing Centre, University of Zagreb; 1999.

[10] Ziha, K, Usage of average uncertainty measures, In: Kalpic D, Hljuz Dobric V, editors. Proceedings of the 22nd International Conference on Information Technology Interfaces; Pula, Croatia. Zagreb: SRCE University Computing Centre, University of Zagreb; 2000.

[11] Hogben, N, Dacunha, N M C, Olliver, G F, Global Wave Statistics, British Maritime Technology Ltd., Feltham, 1986.