



HAL
open science

Classification of probability densities on the basis of Pearson's curves with application to coronal heating simulations

O. Podladchikova, B. Lefebvre, V. Krasnoselskikh, V. Podladchikov

► **To cite this version:**

O. Podladchikova, B. Lefebvre, V. Krasnoselskikh, V. Podladchikov. Classification of probability densities on the basis of Pearson's curves with application to coronal heating simulations. *Nonlinear Processes in Geophysics*, 2003, 10 (4/5), pp.323-333. 10.5194/npg-10-323-2003 . hal-00331066

HAL Id: hal-00331066

<https://hal.science/hal-00331066>

Submitted on 18 Jun 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

Classification of probability densities on the basis of Pearson's curves with application to coronal heating simulations

O. Podladchikova¹, B. Lefebvre², V. Krasnoselskikh², and V. Podladchikov³

¹Max-Planck-Institut für Aeronomie, Max-Planck-Str. 2, D-37191 Katlenburg-Lindau, Germany

²LPCE, CNRS UMR 6115 & Université d'Orléans, 3A av. de la Recherche Scientifique, F-45071 Orléans, France

³Kiev Polytechnic Institute, Department of Applied System Analysis, av. Pobedy 37, Kiev 03056, Ukraine

Received: 26 August 2002 – Revised: 17 December 2002 – Accepted: 18 December 2002

Abstract. An important task for the problem of coronal heating is to produce reliable evaluation of the statistical properties of energy release and eruptive events such as micro- and nanoflares in the solar corona. Different types of distributions for the peak flux, peak count rate measurements, pixel intensities, total energy flux or emission measures increases or waiting times have appeared in the literature. This raises the question of a precise evaluation and classification of such distributions. For this purpose, we use the method proposed by K. Pearson at the beginning of the last century, based on the relationship between the first 4 moments of the distribution. Pearson's technique encompasses and classifies a broad range of distributions, including some of those which have appeared in the literature about coronal heating. This technique is successfully applied to simulated data from the model of Krasnoselskikh et al. (2002). It allows to provide successful fits to the empirical distributions of the dissipated energy, and to classify them as a function of model parameters such as dissipation mechanisms and threshold.

1 Introduction

The precise knowledge of the statistical properties of micro- and nanoflares is a step forward in solving the solar corona heating puzzle.

Due to the central limit theorem, Gaussian distributions are widespread in Nature. Many unusual properties can be described as small deviation from the Gaussian such as rare events modelled by algebraically decaying tails, or power-laws. In solar physics, it is common to approximate the Probability Density Function (PDF) of rare and energetic flares (including micro-flares) by a mixture of power-laws with different indices, which suggests scale invariance of the observed phenomena down to more frequent and less energetic events such as nanoflares. However, it is notoriously

difficult to fit a distribution by a power-law, as for example it often cannot be easily distinguished from an exponential at large values of the argument.

Assuming that flares, microflares and less energetic phenomena are distributed according to a (unique) power-law, the crucial factor that determines the characteristics of the heating process and its scales is related to the index of the power-law. If this index is larger than -2 then flares provide more heating than nanoflares (Hudson, 1991), but not enough to heat the corona. Conversely, an index smaller than -2 indicate that efficient coronal heating may occur at small scales, thanks to an excess in the occurrence rate of small-scale events as suggested by Parker (1988).

For regular flares (generally located in active regions), the observations have mostly been analyzed by peak flux or peak count rate measurements. The results are power-laws, $f(E) \sim E^{-\alpha}$, distributions with indexes α around 1.8 (Dattlowe et al., 1974; Lin et al., 1984; Dennis, 1985; Crosby et al., 1993, 1998; Aschwanden et al., 2001). The later authors also present a substantial investigation of the total energy in the flare electrons observed in hard X-Ray bremsstrahlung, finding $\alpha = 1.5 \pm 0.02$. Soft X-Ray peak fluxes of regular flares have a power-law index between 1.84 (Hudson, 1969) and 1.75 (Drake, 1971). The later authors reports an exponent of the fluence (time-integrated flux, proportional to the total radiated energy) of 1.44. Shimizu (2001) finds a power-law index between 1.5 and 1.6 for flare thermal energy inputs larger than 10^{27} ergs, as determined from soft X-Ray brightenings in active regions. He estimates that the energy supplied by these small flares is at most 20% of the required amount to heat the corona in active regions. The relatively flat distribution ($\alpha < 2$) suggests that the flare below the sensitivity limit cannot be responsible for the rest of the energy input. The very same microflares can be more easily detected in the quiet corona. Soft X-ray observations (Benz et al., 1997) and EUV observations (Harrison, 1997) have revealed enhanced emission and thus intense heating above the magnetic network. A similar phenomenon that forms small X-ray jets at the limb was reported by Koutchmy et al. (1997).

It was found that the number of observed events increases with the sensitivity. It is worth noting that observations of high-temperature iron lines in EUV originate directly from the corona, and allow the derivation of the coronal emission measure with less noise than previous soft X-Ray observations. In addition to these various power-laws with different exponents, non-power-law distributions were found for the statistics of the whole EUV emissions. A χ^2 -distribution (quasi-Gaussian) with a power-law tail of index about -5 for intensities of pixels measured by SOHO/EIT was found (Aletti et al., 2000), as well as the log-normal distribution or a superposition of lognormal and Gaussian distributions for the EUV emission lines in the quiet sun (Pauluhn et al., 2000).

An important result that supports the hypothesis of Parker was reported by Krucker and Benz (1998) who have found from Yohkoh/SXT observations that dissipated energy by nanoflares has a power-law PDF in the range 10^{24} – 10^{26} ergs with the index -2.59 . This result was confirmed by Parnell and Jupp (2000), who estimated the index to be between -2 and -2.1 making use of the data of TRACE. However, Aschwanden et al. (2000) have found significantly different value of this index (-1.80), suggesting that there may not be enough energy in microflares or nanoflares to heat the corona. Differences may arise for several reasons, such as the ways of selecting events, instrumental sensitivity and the models used to translate the emission measures into heat (Benz and Krucker, 2002; Aschwanden and Parnell, 2002), and may as well reveal difficulties with the statistics.

It is also worth mentioning that Benz and Krucker (1998) and Berghmans et al. (1998) have noticed that the heating events occur not only on the magnetic network boundaries but also in the cell interiors. They pointed out that the statistical properties of the faint events that occur in the intra-cell regions of the quiet corona manifest quite small deviations from Gaussian distributions, while supposed nanoflares that are associated with the network boundaries have more prominent enhancements and stronger deviations (see also Harra et al. (2000)).

All the observations show how important it is to provide characterizations of these various distributions as unambiguous as possible. Providing a reliable approximating distribution is essential for the problem of forecasting and comparison with theories and models, and moreover it can be equally important to be able to classify distributions for instance according to the physical process of energy radiation. For this purpose we use a powerful method proposed by K. Pearson at the beginning of the twentieth century. Pearson proposed a classification of distributions according to their first four moments, each class corresponding to well-known distributions. From each class an analytical PDF can be extracted whose agreement with the empirical distribution can be checked by a standard goodness-of-fit test. Although this method is not widely used in the Solar physics community, we propose that it could be useful for the characterization of peak-flux emission measures or energy release and various other problems.

In the present paper, we provide a detailed description of

Pearson's method and apply it to simulated data from a lattice model proposed in Krasnoselskikh et al. (2002) to be described in section 3. The model allows for various sources and magnetic field as well as various magnetic energy dissipation mechanisms. The way how they affect the probability distribution of dissipated energy is studied by means of Pearson's method.

2 Pearson curves for the approximation of statistical distributions

A study of empirical distributions and their fit by theoretical ones should fulfill the following conditions (Hahn and Shapiro, 1967):

- *Objectivity.* A freehand “fit” to the same data may differ from person to person, and it is particularly uneasy to safely estimate power-laws. The use of empirical distributions eliminates such arbitrariness.
- *Automatization.* The data analysis should be as much automatized as possible, in order to allow for fast and economical analysis of large data sets.
- The results of the test should be put in a form as compact and unambiguous as possible, like a small set of parameters.

In accordance with the above criterion, Pearson's distributions open an opportunity to carry out in most cases a unique choice of probability distributions and a classification.

Brief reviews of Pearson's curves are given e.g. in (Hahn and Shapiro, 1967; Kendall, 1958; Wadsworth, 1998). Formulas for each family of them are given in Elderton (1957); Tikhonov (1982). Moreover, a method for computing specific Pearson percentiles is also described in Davis and Stephens (1983). However the problems of classification can not be solved on the basis of an estimation of percentiles only (Johnson et al., 1963; Pearson, 1954). These reviews being relatively condensed, we give a detailed description of the method for reader's convenience.

2.1 Pearson distributions

Pearson distributions are defined by the differential equation

$$\frac{dp(x)}{dx} = \frac{x - a}{b_0 + b_1x + b_2x^2} p(x), \quad (1)$$

where a and b_i are constant parameters of the distribution. Depending the value of these parameters, 12 types of curves can be obtained, which include famous distributions such as the Normal, γ , β , χ^2 and Student which all satisfy Eq. (1) and are special cases of Pearson curves.

Using general properties of probabilities densities, we shall derive expressions for the constants of Eq. (1) for unimodal distributions ($dp/dx = 0$ at a unique $x = a$, which is called the mode).

Eq. (1) can be written as follows

$$x^n (b_0 + b_1x + b_2x^2) \frac{dp(x)}{dx} = x^n (x - a) p(x). \quad (2)$$

Integrating by parts both sides of Eq. 2, and assuming that the distribution decays fast enough at large x

$$\lim_{|x| \rightarrow \infty} x^{n+2} p(x) \rightarrow 0,$$

we get the following recurrence between the moments

$$am_n - nb_0m_{n-1} - (n + 1) b_1m_n - (n + 2) b_2m_{n+1} = m_{n+1} \quad (3)$$

where m_n is the n -th order moment.

For $n = 0, 1, 2, 3$, with $m_{-1} = 0$ and taking into account that $m_0 = 1$, Eq. (3) yields

$$\begin{aligned} -a + b_1 &= 0 \\ b_0 + 3b_2\mu_2 &= -\mu_2, \\ -a\mu_2 + 3b_1\mu_2 + 4b_2\mu_3 &= -\mu_3, \\ -a\mu_3 + 3b_0\mu_2 + 4b_1\mu_3 + 5b_2\mu_4 &= -\mu_4, \end{aligned}$$

where μ_n are the central moments of the distribution. (Assuming without loss of generality that the distribution is centered, $m_1 = 0$, then $\mu_n = m_n$.) From this set of equations one can get

$$a = b_1, b_0 = c_0/d, b_1 = c_1/d, b_2 = c_2/d \quad (4)$$

where

$$\begin{aligned} c_0 &= -\mu_2 (4\mu_2\mu_4 - 3\mu_3^2), \\ c_1 &= -\text{sgn}\mu_3 (\mu_4 + 3\mu_2^2), \\ c_2 &= -2\mu_2\mu_4 + 6\mu_2^3 + 3\mu_3^2, \\ d &= 10\mu_2\mu_4 - 18\mu_2^3 - 12\mu_3^2. \end{aligned}$$

Hence, since the coefficients a, b_0, b_1, b_2 in Eq. (1) can be expressed in terms of m_1, μ_2, μ_3, μ_4 only, the Pearson distributions are entirely determined by their first four moments (provided they exist).

Furthermore, analytical expression of Pearson distributions can be found from Eq. (1). For centered distributions ($a = b_1$), the solution of Eq. (1) reads

$$p(x) = Ce^{\varphi(x)}, \quad (5)$$

where

$$\varphi(x) = \int_0^x \frac{s - b_1}{b_0 + b_1s + b_2s^2} ds. \quad (6)$$

Features of $\varphi(x)$ strongly differ depending on the roots of $b_0 + b_1s + b_2s^2 = 0$, which are

$$s_{1,2} = -\frac{b_1}{2b_2} \left(1 \pm \sqrt{1 - \frac{1}{k}} \right), \quad k = \frac{b_1^2}{4b_0b_2}. \quad (7)$$

Depending on k , several cases are distinguished:

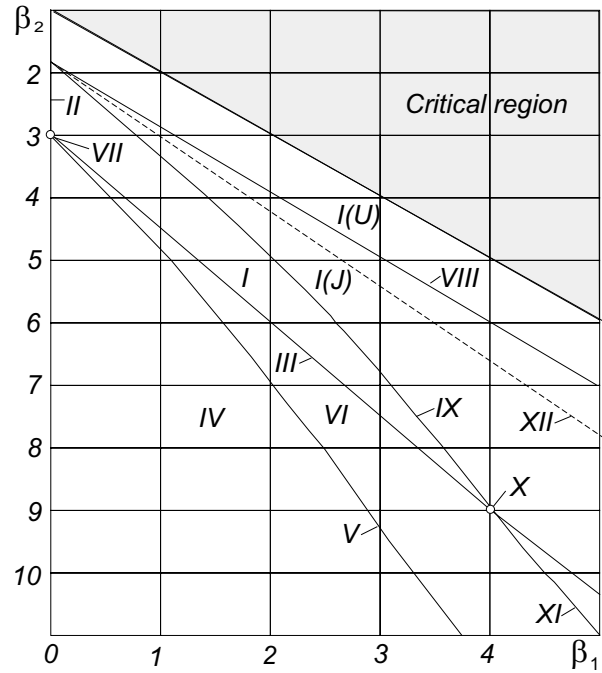


Fig. 1. The diagram of different distributions of Pearson curves family. Here $\beta_1 = \text{skewness}^2$, $\beta_2 = \text{kurtosis} + 3$. The β_2 -axis has been reversed for convenience. The shaded area is forbidden. A detailed presentation of the different classes is given in the Table 1.

- If $k < 0$, roots are real and have different signs. This case corresponds to type-I distribution in Pearson's classification, also called β -distribution. The distribution is defined in $]s_1, s_2[$ (assuming $s_1 < s_2$), and p vanishes outside of this interval.
- If $k > 1$, roots are real and have identical signs (type-VI or β -distribution of the 2nd kind). Depending on this sign, the distribution is non-zero on $] - \infty, s_1[$, or $]s_2, \infty[$.
- If $0 < k < 1$, roots are complex (type-IV distribution). Then p is defined on $] - \infty, \infty[$.

Further distinctions were proposed by Pearson, taking into account particular distributions and boundaries between classes. The normal (type VII) and exponential (type X) distributions are examples of such distributions that fall at the boundary between classes. The total number of classes devised by Pearson is 12. They are summarized together with their analytical form in Table 1.

Since Pearson curves are determined by their first 4 moments, and considering centered distributions, Pearson's classification can be summarized by a diagram (fig. 1) in the (β_1, β_2) plane, where β_1 and β_2 are dimensionless parameters defined by

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3}, \quad \beta_2 = \frac{\mu_4}{\mu_2^2}.$$

Table 1. Pearson distributions

Type	Parameters	Probability density
I β -distribution	$k < 0,$ $x_1, x_2 \in \mathbb{R},$ $x_1 x_2 < 0$	$\begin{cases} p(x) = c x - x_1 ^g x - x_2 ^h \\ or \\ p(x) = cx^{\mu_1 - 1} (1 - x)^{\mu_2 - 1} \end{cases}, x \in [x_1, x_2]$
II β -distribution	$k = \pm\infty$ $x_1 = x_2$ $g = h$	$p(x) = ce^{x/b_1} x + b_0/b_1 ^{1-b_0b_1^{-2}}$
III γ -distribution	$k = \pm\infty$ $b_2 = 0, b_0 = -\mu_2$	if $b_1 = -2, p(x)$ is the χ^2 distribution $\begin{cases} p(x) = c (x^2 + r^2)^{1/2b_2} e^{-\frac{x}{rb_2} \arctan \frac{x}{r}} \\ or \\ p(x) = c \left(\frac{b_0}{b_2}\right)^{1/2b_2} \left(1 + \frac{x^2}{b_0/b_2}\right)^{1/2b_2} \\ or \\ p(x) = \frac{1}{\sqrt{\pi n}} \frac{\Gamma((n+1)/n)}{\Gamma(n/2)} \left(1 + \frac{x^2}{n}\right)^{-(n+1)/2} \end{cases}$
IV	$0 < k < 1$ $x_1, x_2 \in \mathbb{C}$	$\begin{cases} p(x) = c x - x_1 ^g x - x_2 ^h \\ or \\ p(x) = \frac{1}{B(p, q)} \frac{x^{p-1}}{(1+x)^{p+q}} \end{cases}$
V	$k = 1$ $x_1 = -\frac{b_1}{2b_2} = x_2 \in \mathbb{R}$	$p(x) = \frac{ \gamma ^{q-1}}{\Gamma(q-1)} \left x + \frac{b_1}{2b_2}\right ^{-q} e^{-\gamma/(x + \frac{b_1}{2b_2})}$
VI β -distribution	$k > 0$ $x_1 x_2 > 0$ $x_1, x_2 \in \mathbb{R}$	$\begin{cases} p(x) = c x - x_1 ^g x - x_2 ^h \\ or \\ p(x) = \frac{1}{ b_1 } e^{\frac{x-b_1}{b_1}} \text{ for } x \geq b_1 \end{cases}$
VII Gaussian	$b_1 = b_2 = 0,$ $\beta_1 = 0, \beta_2 = 3.$	$p(x) = \frac{1}{\sqrt{-2\pi b_0}} e^{-\frac{x^2}{-2b_0}}$
VIII	idem I, with $g = 0, -1 < h < 0,$ where $h = 0, -1 < g < 0.$	$p(x) = \frac{(x_2 - x_1)^{-h-1}}{B(1, h+1)} (x_2 - x)^h, x \in [x_1, x_2]$
IX	idem I, with $g = 0, h > 0,$ where $h = 0, g > 0$	
X exponential	idem III or IX, with $b_0/b_1^2 = -1$ $\beta_1 = 4, \beta_2 = 9$	
XI	idem VI, with $h = 0$ where $g = 0$	
XII		$p(x) = c \left(\frac{c_1+x}{c_2-x}\right)^m$

Certain distributions are represented by a single point on such a diagram, such as the normal at (0,3), the exponential at (9,4), or uniform distribution at (0,9/5). Other distributions are represented by curves, such as Student- t (of which the normal is a particular case, and which belong to type IV) or the γ -distribution (type III). Finally, certain more general classes such as β -distributions (type I) occupy finite area on the diagram. In particular, power-laws belong to the later class. However, certain distributions, such as the Cauchy distribution (which formally belongs to type IV), cannot be represented on the diagram, if at least one of the first four moment does not exist.

It is also worth noting that probability distributions are confined to a certain area in the (β_1, β_2) plane. Indeed,

for any centered random variable ξ , making use of Cauchy-Bunyakovsky inequality we have

$$\begin{aligned} \beta_1 &= \left\langle \frac{\xi}{\mu_2^{1/2}} \left(\frac{\xi^2}{\mu_2} - 1 \right) \right\rangle^2 \leq \frac{\langle \xi^2 \rangle}{\mu_2} \left\langle \left(\frac{\xi^2}{\mu_2} - 1 \right)^2 \right\rangle = \\ &= \frac{\mu_4}{\mu_2^2} - 1 = \beta_2 - 1 \end{aligned}$$

Hence, for any probability distribution the inequality

$$\beta_2 \geq \beta_1 + 1$$

must be true. The area which does not satisfy this inequality is shaded in gray on Fig. 1.

2.2 Approximation technique by means of Pearson curves.

The parametric technique of approximation of observed distributions determined from time series by Pearson curves can be reduced to the following stages:

1. From the time series, one computes the first 4 moments and (β_1, β_2) , which determines the distribution's class.
2. According to expressions (4) one takes the values of experimental moments as the moments of chosen theoretical distribution.
3. Determining the integral (6) and substituting in (5) one can find the required distribution.
4. Validating the results by comparing the obtained PDF with the statistical one.

It is worth noting that the estimation of the moments, and thus the whole procedure, uses all the samples at hand, and thus uses more information from the initial data than the simplest techniques which consist in fitting an histogram to obtain an empirical Pearson distribution.

To estimate if the chosen theoretical probability density agrees well with the results of observations (step 4), one can use the best-fit criteria. This is an important stage, particularly since errors may arise in the estimation of the moments. The most used (and strictest) criteria is probably Pearson's χ^2 (Bendat and Piersol, 1986).

3 A statistical model of magnetic energy dissipation in the corona

The method shall now be applied to simulations of a simple lattice model for small scale heating in the lower quiet corona (Krasnoselskikh et al., 2002). The output of the model is a time series of dissipated magnetic energy transformed into heat, whose PDF was found to depend on the dissipation mechanism considered, as well as on the way the magnetic energy (magnetic flux emergence) is deposited into the system. The goal here is to classify these distributions and precisely characterize their shapes.

The model consists of a 2D grid representing a cross-section of the low solar atmosphere, where a magnetic field perpendicular to the grid is distributed in each cell. At each time step, magnetic energy is deposited into the system. In each cell, a quantity δB is added to the previous value, where δB may be a random variable equiprobably chosen in $\{-1, 0, 1\}$ (other types of sources were considered by Krasnoselskikh et al. (2002); Podladchikova et al. (2001)).

The magnetic energy dissipation and transformation into heat depends on the currents. Currents are computed as $\mathbf{j} = \nabla \times \mathbf{B}$, and circulate on the borders between cells. Two dissipation mechanisms relevant for coronal physics are considered:

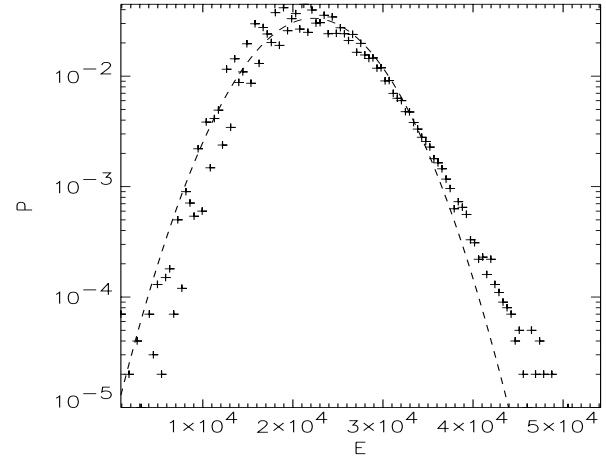


Fig. 2. PDF of dissipated energy at $j_{\max} = 32$, reconnection. Dotted line shows the Gaussian distribution with the same mean and variance as statistical one.

- anomalous resistivity, which arises when the current exceeds a threshold for micro-instabilities such as modified Buneman. Here, currents are dissipated whenever they exceed a given threshold $j > j_{\max}$
- reconnection, which here is supposed to occur when a configuration similar to a X-point occurs (B has opposite signs in neighboring cells) and when the current exceeds the threshold j_{\max} .

The energy dissipated over the grid is computed as the sum of the dissipated j^2 . The model is described in more detail in (Krasnoselskikh et al., 2002; Podladchikova et al., 1999).

4 Estimation of dissipated magnetic energy distribution

A detailed application of Pearson's technique is given for simulations of uniformly random sources of magnetic field on a 128×128 grid during $t = 10^5$. The currents dissipate by magnetic reconnection when they achieve the critical value $j_{\max} = 32$.

The dissipated energy is bounded by $E_{\min} = 1024$ and $E_{\max} = 5.416 \cdot 10^4$. This interval is broken into 120 equal intervals (classes). The histogram of dissipated energy, in the stationary regime, is shown on Fig. 2.

The procedure is decomposed into the following steps:

- *Moments.* The average is $m_1 = 2.225 \cdot 10^4$, the variance $\mu_2 = 2.901 \cdot 10^7$, the skewness $\gamma_1 = 0.28$ and the kurtosis $\gamma_2 = 0.37$. From this we get

$$\beta_1 = \gamma_1^2 = 0.08, \quad \beta_2 = \gamma_2 + 3 = 3.37$$

- *Distribution type.* The character of the roots (7) is determined by the value of $k = 0.13$. Since $0 < k < 1$, the roots (7) are complex, which corresponds to type-IV in Pearson's classification.

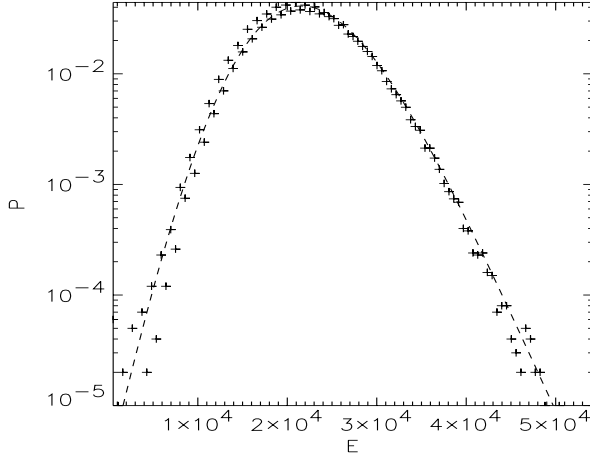


Fig. 3. Probability density of dissipated energy at $j_{\max} = 32$, reconnection. The dotted line shows the distribution found by Pearson's method. The fit is accepted with a significance value $\alpha = 0.29$ according to χ^2 -criteria.

A special case of this distribution at $\beta_1 = 0$, $\beta_2 \leq 3$ is the Student distribution, and $\beta_2 = 3$ corresponds to the normal distribution law (type VII).

- *Distribution's parameters.* Using the Eq. (4), we shall determine the distribution's parameters

$$b_0 = -2.6 \times 10^7, \quad b_1 = -671.7, \quad b_2 = -0.03$$

It is thus quite different from the normal distribution, which corresponds to $b_1 = b_2 = 0$.

- *The distribution.* For such parameters, the function φ defined by Eq. (6) can be explicitly written as

$$\varphi\left(x - \frac{b_1}{2b_2}\right) = \frac{1}{2b_2} \ln(x^2 + A^2) - \frac{B}{Ab_2} \arctan \frac{x}{A} + c$$

with

$$B = b_1 \left(1 + \frac{1}{2b_2}\right), \quad A^2 = \frac{b_0}{b_2} - \frac{b_1^2}{4b_2^2}.$$

Hence the probability density according to expression (5) reads

$$p\left(x - \frac{b_1}{2b_2}\right) = C \left[\left(\frac{x}{A}\right)^2 + 1 \right]^{\frac{1}{2b_2}} e^{-\frac{B}{Ab_2} \arctan \frac{x}{A}}$$

for $-\infty < x < \infty$, with $A = 2.6 \times 10^4$, $B = 9.5 \times 10^3$ and $b_2 = -0.03$. Given these coefficients,

$$p(x) = 1.1 \times 10^{-5} \frac{e^{11 \arctan f(x)}}{(f(x)^2 + 1)^{15}}$$

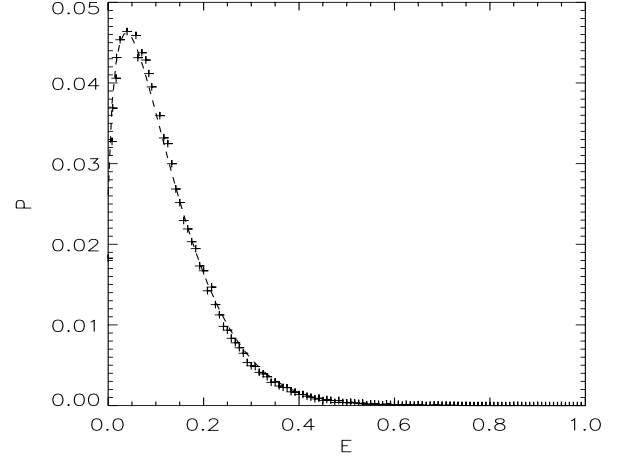


Fig. 4. Tail of the PDF of Fig. 3. Energies below $E_0 \approx 3 \times 10^4$ were cut, and data was rescaled between 0 and 1. The β -distribution found by Pearson's method is shown by a dotted line.

where $f(x) = \frac{x - 2.22 \times 10^4}{2.6 \times 10^4} + 0.39$, and C is found from the normalization condition. Let us notice that this distribution is centered, and thus shifted with respect to the original one by m_1 .

A comparison between the experimental histogram and the obtained Pearson curve is shown on Fig. 3. The agreement seems satisfactory both in the core of the distribution (area of high probabilities) and in its tails (in the areas of small probabilities). Indeed, Pearson's χ^2 -criteria was applied to check the accuracy of this fit. The hypothesis that Pearson's distribution fits the experimental one is accepted with a very high level of the significance $\alpha = 0.29$, at $r = 100$ (while for a Gaussian hypothesis, $\alpha = 10^{-9}$).

5 Probability of high-energy events

As we have previously seen, distribution functions of the dissipated energy are often rather close to the Gaussian, the most significant difference being in the tail of the distribution which correspond to high energy events. Another reason to study high-energy events is that experiments cannot resolve micro- or nanoflares down to arbitrary small energy, and thus introduce some sort of low-energy cut-off corresponding to the minimal energy and scale of the observed events.

To study high-energy events, we have simply filtered out energy releases smaller than a minimal E_0 , and replaced the time series of the dissipated energy values smaller than E_0 by white noise. Typically, E_0 can be taken of the order of the mean plus standard deviation. Then data are conveniently normalized so that E varies between 0 and 1.

The histogram of dissipated energy after such a filtering is presented on Fig. 4. The time series is the same as the one analyzed in the previous section (the grid is 128×128 , $t = 10^5$, currents dissipate by reconnection and $j_{\max} = 32$). The maximum value of energy is $E_{\max} = 5.4 \cdot 10^4$ and we

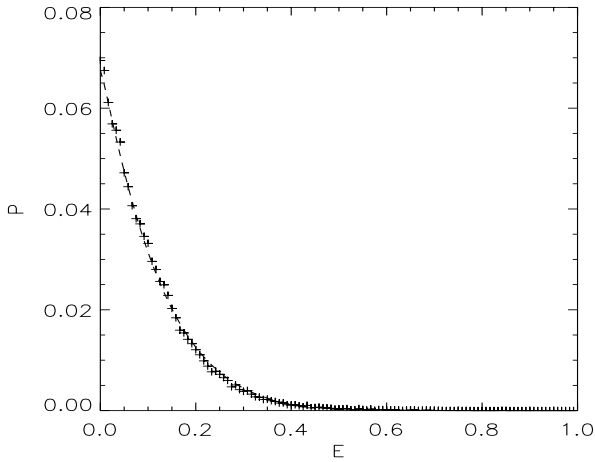


Fig. 5. Tail of the dissipated energy PDF, for the same parameters as in Fig. 4 but with a sharp cut-off at low values. The obtained β -distribution by Pearson’s method is shown by a dotted line.

take $E_0 = 2.97 \cdot 10^4$. The range of energy values $E_{\max} - E_0$ is broken into 120 equal intervals.

Although the distribution seems to follow a power-law, the found significance value $\alpha = 1 \times 10^{-12}$ from χ^2 -criteria rejects this hypothesis.

5.1 β -distributions

Therefore the choice of probability density was made from the family of Pearson curves. Here we get $b_0 = -5.3 \times 10^6$, $b_1 = -1678$, $b_2 = 0.008$ and

$$\beta_1 = 2.05, \quad \beta_2 = 6.0,$$

and according to formula (7) $k = -16.23$. As $k < 0$, this corresponds to a β -distribution which is non-zero in the interval $[x_1, x_2]$.

Let us write the integrand in (6) as

$$\begin{aligned} \frac{x - b_1}{b_0 + b_1x + b_2x^2} &= \frac{x - b_1}{b_2(x - x_1)(x - x_2)} \\ &= \frac{g}{x - x_1} + \frac{h}{x - x_2}, \end{aligned}$$

where

$$g = \frac{b_1 - x_1}{b_2(x_2 - x_1)}, \quad h = \frac{x_2 - b_1}{b_2(x_2 - x_1)}. \tag{8}$$

Then

$$\begin{aligned} \varphi(x) &= \int_0^x \left(\frac{g}{s - x_1} + \frac{h}{s - x_2} \right) ds \\ &= g \ln|x - x_1| + h \ln|x - x_2|. \end{aligned}$$

Since $x_1 \leq x \leq x_2$, then

$$p(x) = C(x - x_1)^g(x_2 - x)^h. \tag{9}$$

As already noted, the power-law is a special case of this distribution.

The normalizing coefficient is found to be

$$C = \frac{1}{(x_2 - x_1)^{g+h+1} \beta(g+1, h+1)}$$

where $\beta(x, y)$ is the β -function defined by

$$\beta(p, q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)} = \int_0^1 t^{p-1}(1-t)^{q-1} dt$$

(Abramowitz and Stegun (1974)).

Writing $z = (x - x_1) / (x_2 - x_1)$, we get a more usual form for this distribution

$$p(z) = \frac{\Gamma(\mu_1 + \mu_2)}{\Gamma(\mu_1)\Gamma(\mu_2)} z^{\mu_1-1} (1-z)^{\mu_2-1},$$

whose moments read

$$\mu_1 = \frac{m_1(m_1 - m_2)}{m_2 - m_1^2}, \quad \mu_2 = \frac{(1 - m_1)(m_1 - m_2)}{m_2 - m_1^2},$$

where m_1 and m_2 are respectively the first and second order moment of $p(x)$.

Following the same procedure as in the previous section, and using the values of the moments, we obtain an approximated distribution density of the following form:

$$p(z) = 29.15 \times z^{0.4} (1 - z)^{9.0}.$$

This PDF is shown in Fig. 4 as a dashed curve, and is seen to coincide with the statistical PDF on the whole interval of energies. It is not a simple power-law, but a slightly more complicated β -distribution. One can notice that the cut-off at low-energies is rather smooth, although rather abrupt, due to the presence of white noise.

If we impose a sharp cut-off at low-energies, by simply cancelling all energies smaller than E_0 , the distribution is also found to belong to type-I (Fig. 5). However, the parameters

$$\beta_1 = 2.85, \quad \beta_2 = 7.14$$

are slightly different, and the found Pearson curve reads

$$p(z) = 9.47 \times z^{0.02} (1 - z)^{8.0}$$

The rather good coincidence of the statistical distribution with the theoretical one is clear on the figure, and was confirmed by the χ^2 test.

5.2 γ -distributions

It is necessary to notice that the β -distributions obtained in the both of cases of cut-off are located near the border of the area on Pearson’s diagram, determined by the condition

$$2\beta_2 - 3\beta_1 - 6 = 0,$$

between the β -I distribution (type I in Pearson’s classification) and the β -II distribution (type VI).

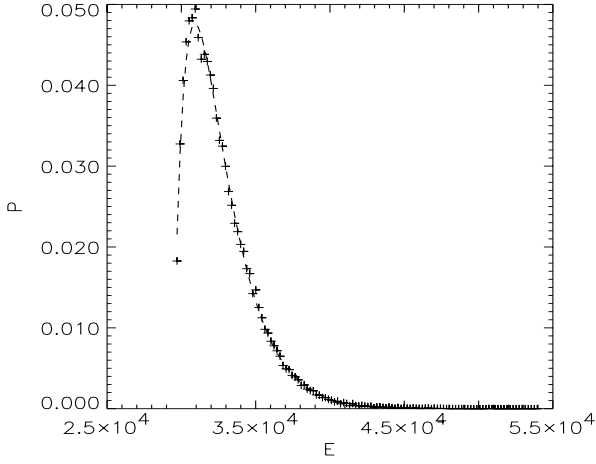


Fig. 6. The same series with the noisy cut-off of low-energy events as on the figure 4, but approximated by the γ -distribution shown by a dotted line, and accepted with a significance value $\alpha = 0.33$ by χ^2 -criteria, for $r = 30$.

Thus, for a noisy cut-off of lower energies

$$2\beta_2 - 3\beta_1 - 6 = 0.15.$$

The points on this border correspond to type III distributions or γ -distribution. Hence, one can accept the hypothesis that statistical distribution of high-energy events can be approximated by such a γ -distribution:

$$p(z) = \frac{1}{-b_1 \cdot \Gamma(\alpha + 1)} \left(\frac{z}{-b_1} \right)^\alpha e^{z/b_1},$$

where $z = x + \frac{b_0}{b_1} - m_1$; $\alpha = -1 - \left(\frac{b_0}{b_1^2} \right)$. Special cases of this distribution are the χ^2 -distribution and the exponential distribution (see the Diagram 1).

Then for the noisy cut-off the distribution looks like (see Fig. 6):

$$p(z) = 6.29 \times 10^{-5} z^{0.30} e^{-z/1840},$$

and for the case of a sharp cut-off at low energies:

$$p(z) = 9.765 \times 10^{-7} z^{0.87} e^{-z/1678}.$$

This distribution is closer to the exponential on the Diagram 1. The approximation of statistical distribution shown on the figures 6 and 7 by the γ -distribution is not less precise than their approximation by a β -distribution. Pearson's χ^2 -criteria gives the significance values $\alpha = 0.33$ and $\alpha = 0.37$ respectively, at $r = 30$.

6 Dependence of the distribution on model parameters

Although a detailed presentation of the method was given, all the procedure can be done rather automatically. This allows to take great advantage of Pearson's method by classifying

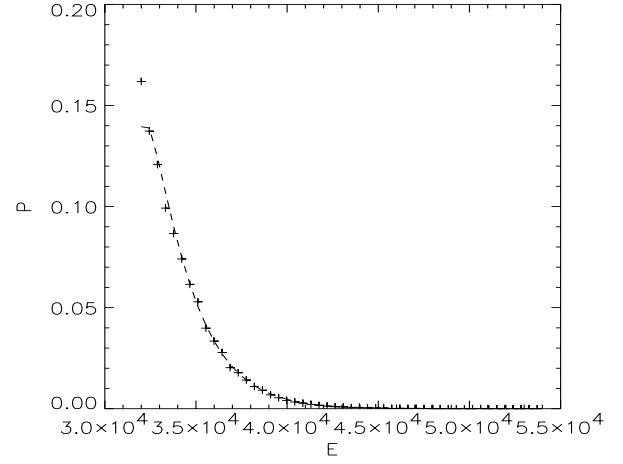


Fig. 7. The same series with a sharp cut-off of low-energy as on the figure 5, but approximated by the γ -distribution shown by a dotted line. It accepted with significance level $\alpha = 0.37$ at $r = 30$ according to χ^2 best-fit criteria.

and studying parameters's evolution of the distribution function as the model's parameters are varied.

Whereas dissipated energy distribution appears to be rather insensitive to the source's properties, it strongly depends on the dissipative processes. Increasing the dissipation threshold j_{\max} , PDFs more and more strongly depart from thermal (Gaussian) distribution. With the growth of j_{\max} supra-thermal tails of distributions become appreciable, and a supra-thermal tail forms at high energies (with different rates for different sources) and visually remind the power law.

More precisely, an approximation of the dissipated energy PDF by Pearson curves method was systematically applied to several simulation runs with different thresholds and dissipation mechanisms. Results can be summarized as follows:

- For any type of source of magnetic field and dissipation mechanism, for small (of the order of 1 or smaller) dissipation threshold j_{\max} the distribution function of dissipated energy is close to normal. The best approximation of the statistical law of distribution is the type-IV distribution in Pearson's classification. This result is validated by goodness of fit criteria. A limit case, a normal law, of this distribution is reached at small thresholds (see a fragment of the Diagram 1 with trajectories of the obtained functions of distributions on Fig. 8).
- With the growth of j_{\max} , the distributions tends to β -distribution for all types of sources. The first β -distributions have a global maximum, while when the threshold is further increased it degenerates to the monotonous decreasing case, the power-law distribution (see Diagram 8). It is important to note that the approach to power-law distribution with the increasing of j_{\max} is faster with reconnection than for anomalous resistivity.

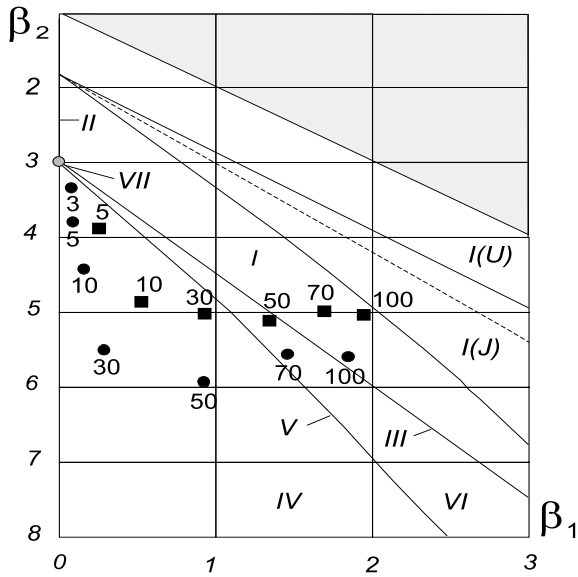


Fig. 8. Evolution of probability distribution of dissipated energy, with various dissipation thresholds, for the grid 200×200 , and reconnection (squares) or anomalous resistivity (circles), in Pearson’s classification. Distribution density evolve from type-IV (nearly Gaussian, point 3 on the diagram), to β -distributions. Both curves approach the power-law (type I(J)) for large current density thresholds, although reconnection processes tend faster toward the power-law.

- The supra-thermal tails of distributions in high energies are approximated by the distributions of the type $p(x) \approx A \cdot x^{-\alpha} e^{-x}$, or γ -distributions. A rather narrow class of points (the straight line on the Diagram 1) corresponds to this distribution, on the border of β -distributions. The contribution of power and exponential co-factors varies depending on parameters.

7 Discussion and conclusion

Pearson’s technique of approximation of data sets by empirical distributions was described and applied to results from a simple model of small-scale coronal heating. The dependence of the statistical properties of the dissipated energy (heat) on the dissipation mechanism of magnetic energy was studied carrying out a series of simulations with various dissipation thresholds depending on currents and various mechanisms of dissipation (anomalous resistivity and local reconnection).

The results of simulations are summarized in Fig. 8. They were found to be independent of grid size for sufficiently large grids. We have found that for any source type of magnetic field and dissipation mechanism, the parameters β_1 and β_2 belong to the range of possible distributions permitted by the Pearson approximation, and that an empirical distribution could always be found for all cases under investigation.

We could follow in a clear and precise way the qualitative evolution, deviation from Gaussianity and shape evolution of the PDF as the parameters (here the instability threshold) change. This allowed to show that for both dissipation criteria, the PDFs followed a similar evolution, crossing the same classes and deviating farther from Gaussianity by developing heavier tails. Moreover, it allowed to show that with the reconnection mechanism the probability densities tend faster toward the power-law. Indeed, for a given threshold, with reconnection the PDF has a higher skewness, and at large thresholds (greater than 30) it also has a higher kurtosis which shows that it develops heavier tails. The formation of a power-law tail was qualitatively explained by the increase of the energy storage time between dissipative events by Krasnoselskikh et al. (2002). As a consequence, events corresponding to high energy releases contribute more strongly to the heating. Finally, Fig. 8 shows that for very large thresholds the statistics of energy release become quite independent from the dissipation mechanism, as both dissipation mechanisms produce distributions which are close to each other (in the (β_1, β_2) plane) and tend to the same class.

Several more general comments may help clarifying the strengths and limitations of the method, as we see them. First, having enough information to find a theoretical distribution is always better than to fit an empirical model to data, although it is rather rarely the case for such complex problems as coronal heating. Second, it should be clear that Pearson curves do not contain all known distributions, not even all those satisfying the (not very restrictive for our problem) technical requirements of unimodality and smoothness. For instance, a different set of empirical curves was proposed by Johnson (Hahn and Shapiro, 1967; Wadsworth, 1998). Also, not all distribution can be characterized by their first four moments only, although in practice higher order moments are very often useless due to high statistical errors. Hence, Pearson curves could also miss several interesting distributions from a purely theoretical point of view. Although it is improbable that there will ever exist a method allowing to find a unique distribution which best fits any given data set without a priori hypotheses, an interesting property of Pearson’s method is that it provides a unique empirical function among Pearson curves for given β_1 and β_2 , the only uncertainty coming from the possible errors on β_1 and β_2 . Moreover, these curves cover a huge range of possibilities. Even in the case of distributions whose first four moments may not be theoretically finite, in practice a finite amount of samples provides finite moments, and we have found in most cases that Pearson’s technique allows to reconstruct correctly the distribution on the interval in which samples are observed (Podladchikova, 2002).

Pearson’s technique also has the advantage to produce analytical expressions for the densities containing small sets of parameters, and to allow a simple graphical representation and classification as in Fig. 8. Therefore results can be presented in a compact form, and can be obtained in a rather automatic way. Having an analytical form can be useful for comparisons with theories, interpolating or performing

Monte-Carlo simulations for example. Moreover, these results are objective, and do not depend on a subjective fitting by eye which pushes to fit many slowly decreasing functions by power-laws. Finally we would like to stress that this technique uses all data points, and thus retains much more information from the original data sets than the simplest methods based on binning data into an histogram which is then fitted by a given curve belonging to one of Pearson's classes.

We expect that the technique shall also be useful (among other problems) to test the hypothesis of coronal heating by small-scale impulsive events, in particular for statistics of EUV brightenings. At this stage of the coronal heating problem, which is mainly to select between different possible mechanisms, a precise and as objective as possible knowledge of the statistical properties of the diverse heating phenomena is essential.

Since the events that are looked for are near the instruments' resolution and sensitivity, their distribution is contaminated by Gaussian instrumental noise and appears as deviations from the Gaussian in the high intensity tail. This complicates a lot the task of selecting events (Benz and Krucker, 2002) and characterizing the shape of their distribution. Pearson's method could provide some help at this stage. Moreover, getting a distribution shape as precise as possible is even more important if, as is probably the case, the impulsive events which contribute the most to the heating are those below the observational capabilities so that some extrapolation of distributions toward smaller energies is required to estimate their contribution, even though such an extrapolation is always rather hazardous. Another application would be to classify the observed distributions, for example above the magnetic network or the cell interiors, or for regions of various activity levels. Application to simulations was a first step before the study of experimental data, which we plan to do in a near future.

Acknowledgements. The authors thank J.Y. Brochot for a careful reading of an early version of the manuscript. Discussions with T. Dudok de Wit are also gratefully acknowledged. OP acknowledges a post-doctoral fellowship from the Max Planck Institut, and BL from the CNES.

References

- Abramowitz, M. and Stegun, I. A.: Handbook of Mathematical Functions, With Formulas, Graphs, and Mathematical Tables, Dover, 1974.
- Aletti, V., Velli, M., Bocchialini, K., Einaudi, G., Georgoulis, M. and Vial, J. C.: Microscale Structures on the Quiet Sun and Coronal Heating, *Astrophys. J.*, 544, 550–557, 2000.
- Aschwanden, M. J. and Parnell, C. E.: Nanoflare Statistics from First Principles: Fractal Geometry and Temperature Synthesis, *Astrophys. J.*, 572, 1048–1071, 2002.
- Aschwanden, M. J., Poland, A. I., and Rabin, D. M.: The New Solar Corona, *Ann. Rev. Astron. Astrophys.*, 39, 175–210, 2001.
- Aschwanden, M. J., Tarbell, T. D., Nightingale, R. W., and Schrijver, C. J.: Time Variability of the “Quiet” Sun Observed with TRACE. II. Physical Parameters, Temperature Evolution, and Energetics of Extreme-Ultraviolet Nanoflares, *Astrophys. J.*, 535, 1047–1065, 2000.
- Bendat, J. S. and Piersol, A. G.: Random Data. Analysis and measurement procedures., Wiley-Interscience publication, 1986.
- Benz, A. O. and Krucker, S.: Heating Events in the Quiet Solar Corona, *Sol. Phys.*, 182, 349–363, 1998.
- Benz, A. O. and Krucker, S.: Energy Distribution of Microevents in the Quiet Solar Corona, *Astrophys. J.*, 568, 413–421, 2002.
- Benz, A. O., Krucker, S., Acton, L. W., and Bastian, T. S.: Fine structure of the X-ray and radio emissions of the quiet solar corona, *Astron. Astrophys.*, 320, 993–1000, 1997.
- Berghmans, D., Clette, F. and Moses, D.: Heating Events in the Quiet Solar Corona, *Astron. Astrophys.*, 336, 1039, 1998.
- Crosby, N. B., Aschwanden, M. J., Dennis, B. R., and Brian, R.: Frequency distributions and correlations of solar X-ray flare parameters, *Sol. Phys.* 143, 275–299, 1993.
- Crosby, N. B., Vilmer, N., Lund, N., and Sunyaev, R.: Deca-keV X-ray observations of solar bursts with WATCH/GRANAT: frequency distributions of burst parameters, *Astron. Astrophys.*, 334, 299–313, 1998.
- Datlowe, D. W., Eclan, M. J., and Hudson, H. S.: OSO-7 observations of solar X-rays in the energy range 10–100 keV, *Sol. Phys.*, 39, 155–174, 1974.
- Davis, C. S. and Stephens, M. A.: Approximate percentage points using Pearson curves, *Applied Statistics*, 32, 322–327, 1983.
- Dennis, B. R.: OSO-7 observations of solar X-rays in the energy range 10–100 keV, Solar hard X-ray bursts, *Sol. Phys.*, 100, 465–490, 1985.
- Drake, J. F.: Characteristics of Soft Solar X-Ray Bursts, *Sol. Phys.*, 16, 152–185, 1971.
- Elderton, W. P.: Frequency curves and Correlations, 4th ed., Cambridge University Press, 1957.
- Hahn, G. J. and Shapiro, S. S.: Statistical models in engineering, Wiley (New-York), 1967.
- Harra, L., Gallagher, P. T., and Phillips, K. J. H: Characteristics of quiet Sun cell and network brightenings, *Astron. Astrophys.*, 362, 371–378, 2000.
- Harrison, R. A.: One Year of CDS: Highlights from Observations using the Coronal Diagnostic Spectrometer on SOHO, in: The Corona and Solar Wind Near Minimum Activity, ESA SP-404, 7, 1997.
- Hudson, H. S.: Solar flares, microflares, nanoflares, and coronal heating, *Solar Physics*, 133, 357–369, 1991.
- Hudson, H. S., Peterson, L. E., and Schwartz, D. A.: The Hard Solar X-Ray Spectrum Observed from the Third Orbiting Solar Observatory, *Astrophys. J.*, 157, 389–416, 1969.
- Johnson, N. L., Nixon, E., and Amos, D. E.: Tables of Percentage Points of Pearson Curves for given $\sqrt{\beta_1\beta_2}$, expressed as Standard Measures, *Biometrika*, 50, 459–471, 1963.
- Kendall M. G. and Stuart A.: The advanced theory of Statistics, vol. I, Hafner Publishing Company (New York), 1958.
- Koutchmy, S., Hara, H., Suematsu, Y., and Readron, K.: SXR Coronal Flashes, *Astron. Astrophys.*, 320, L33, 1997.
- Krasnoselskikh, V., Podladchikova, O., Lefebvre, B., and Vilmer, N.: Quiet Sun coronal heating: a statistical model, *Astron. Astrophys.*, 382, 699–712, 2002.
- Krucker, S. and Benz, A. O.: Energy Distribution of Heating Processes in the Quiet Solar Corona, *Astrophys. J.*, 501, L213, 1998.
- Lin, R. P., Schwartz, R. A., Kane, S. R., Pelling, R. M., and Hurley, K. C.: Solar hard X-ray microflares, *Astrophys. J.*, 283, 421–425, 1984.
- Parker, E. N.: Nanoflares and the solar X-ray corona, *Astrophys. J.*,

- 330, 474–479, 1988.
- Parnell, C. E. and Jupp, P. E.: Statistical Analysis of the Energy Distribution of Nanoflares in the Quiet Sun, *Astrophys. J.*, 529, 554–569, 2000.
- Pauluhn, A., Solanki, S. K., Rüedi, I., Landi, E., and Schühle, U.: Statistics of quiet Sun extreme ultraviolet intensities, *Astron. Astrophys.*, 362, 737–745, 2000.
- Pearson, E. S. and Hartley, H. O.: *Biometrika Tables for Statisticians*, Vol.1, Cambridge University Press, Cambridge, 1954.
- Podladchikova, O.: PhD dissertation, Orléans University and Kiev National University, 2002.
- Podladchikova, O., Krasnoselskikh, V., and Lefebvre, B.: in: *Magnetic fields and Solar Processes*, ESA SP-448, Vol. 1, 553–559, 1999.
- Podladchikova, O., Dudok de Wit, T., Krasnoselskikh, V., and Lefebvre, B.: Quiet Sun coronal heating: analyzing large scale magnetic structures driven by different small-scale, uniform sources *Astron. Astrophys.*, 383, 713–721, 2002.
- Shimizu, T.: Energetics and Occurrence Rate of Active-Region Transient Brightenings and Implications for the Heating of the Active-Region Corona, *PASJ*, 47, 251–263, 2001.
- Tikhonov, V. I.: *Statistical radio engineering*, Radio and Communication, Moscow, 1982.
- Wadsworth, H. M. (Ed): *Handbook of Statistical Methods for Engineers and Scientists*, 2nd edition, McGraw-Hill Professional Publishing, 1998.