

Experience in Spatial Verification of Precipitation Radar Nowcasting

A.V. Muraviev, A.Yu. Bundel

Introduction

This study presents the results of the object-oriented verification of the precipitation radar nowcasting system of the Hydrometeorological Center of Russia. The verification period is May-September 2017 and November 2017 – March 2018. The focus is on the ability of the system to reproduce contiguous precipitation areas (“objects”) with an area exceeding a certain threshold. The generalized three-parameter Pareto distribution is used, which enables evaluation of the area sizes in the distribution tails by the shape parameter of the distribution.

The paper [Muraviev 2022b] discusses the problems of constructing correct samples for verification and the statistical properties of these samples based on the accumulated datasets of observations and forecasts. Here, the main attention is paid to the precipitation objects in the sense of object-oriented verification of weather elements [Davis 2006 a, b, Bundel 2021]. The objects are determined in the following way: 1) spatial smoothing of the field with a certain smoothing radius; 2) identification of connected areas using the precipitation intensity threshold - in fact, *objects*; 3) selection of objects with an area exceeding a given threshold value.

The precipitation intensity was derived from radar reflectivity based on the methods of the Central Aerological Observatory (CAO) [Temporary guidelines 2017].

Since the precipitation fields belong to time series with a 10-minute step, the independence of sample elements can be violated due to serial correlation. To obtain conditionally independent objects of maximum area, the concept of a continuous situation is applied, under the assumption that particular situations in the synoptic sense can be considered as physically partially independent [Muraviev 2022b].

The area and the object are defined in precipitation fields using the isoline of the same value, with the difference that the area belongs to the original field, and the object belongs to the same field, but after spatial averaging. It is possible that the area exists, but the corresponding object is not defined due to the smoothing procedure, which reduces the initial intensity of precipitation. From a meteorological point of view, the area of the object corresponds to the spatial scales of mesoprocesses. It could be more natural to switch to a linear scale by taking the square root of the number of points in the object [Davis 2006a], but objects in precipitation fields have not only ellipsoid or square shapes, but also elongated stripes and even connected lines [Isaev 2001, Lenskaya 2006]. Thus, the number of points seems to be more informative estimate of the object's scale in this work.

We used four methods for estimating the parameters of generalized distributions of extremes and threshold exceedances: 1) maximum likelihood, 2) generalized maximum likelihood, 3) L-moments, and 4) Bayesian estimation using Monte Carlo procedures. Confidence intervals (CI) of parameter estimates of the generalized Pareto distribution calculated from sets of radar observations and nowcasts based on the STEPS model are used to integrally assess the quality of the statistical nowcasting model according to the simplest principle: the greater the intersection of confidence intervals for estimates of the same parameters in observations and forecasts, the higher quality of forecasts, in terms of the model ability to reproduce the observed distribution of vast precipitation areas.

The resulting tables collect CI intersection estimates for network radars for the warm and cold periods of 2017-2018, and for several area thresholds. Particular attention is paid to the quality of forecasting the heavy tails of the distribution. In the conclusions, considerations are formulated about the information content of such quality assessments and about the possibilities of their generalization.

1. Construction of a sample of objects with large and extreme areas

The correct sample formation is considered in detail in [Muraviev 2022b]. The selection of objects in a given field is performed using the FeatureFinder function of the SpatialVx library [Gilleland 2020] based on three parameters: 1) the radius of spatial averaging, 2) the identification threshold, 3) the minimum area - by the number of connected, or adjacent points of the object. The following parameter values were tested: averaging radii - 5 and 9 points; object identification threshold of 1 mm/h; the minimum sizes of objects (according to the number of grid points in the corresponding area) are 0, 625, 900, 1225 and 1600 points. In areal units, the minimum dimensions correspond to the squares 50×50 , 60×60 , 70×70 and 80×80 [km $\times$ km]. We do not take into account the problem of the identified object location at the radar coverage boundary at the present stage, and cannot evaluate to what extent the objects surpassing the boundary influence the results. In the most important cases, the visual analysis can be used. When passing to radar composite data covering the whole territory of European Russia, the influence of the boundaries should considerably lessen.

To identify an object, it is recommended to use two methods for determining the threshold: as a fixed value or as a distribution quantile [Davis 1, 2]. Both approaches are complementary, each has advantages and disadvantages. In our work, we use a fixed threshold due to the large scatter in precipitation fields for different radars and in different test periods [Muravyov 2019].

The distribution of precipitation object sizes for the fields of radar observations and forecasts was analyzed over the coverage areas of nine DMRL-S radars and in 10-minute time step. The distribution of object sizes by quartiles and with an average value (mean) is presented in Table 2. A somewhat unexpected fact is that in all radar view areas the smallest object is one point, or one cell.

Table 2. Statistical Characteristics of the Size Distribution of Verification Objects in the Observation Fields of different Radars (RAKU, RATL, etc.) in the Central Federal District of Russia.

	min	q 25	median	mean	q 7 5	max
RAKU 15403 / 7384						
warm	one	125	317	1054	1133	17989
cold	one	76	262	1017	1080	14228
RATL 13269 / 4511						
warm	one	116	304	985	1053	16682
cold	one	86	288	1000	1098	22158
RAVN 14094 / 1502						
warm	one	122	316	848	916	16855
cold	one	71	203	511	523	5654
RAVO 21944 / 5689						
warm	one	92	257	749	731	14820
cold	one	65	184	567	514	10560
RUDB 12022 / 4120						
warm	one	134	361	1037	1164	19383
cold	one	83	270	996	980	12528
RUDK 20455 / 3866						
warm	one	104	290	924	970	14443
cold	one	66	180	557	505	11824
RUDL 20479 / 4246						
warm	one	107	298	936	1007	19812
cold	one	72	233	905	841	27128
RUDN 21215 / 5281						
warm	one	113	301	971	1004	17049
cold	one	62	193	704	614	14079
RUWJ 22991 / 2445						
warm	one	95	267	953	895	19551
cold	one	54	172	527	478	16530

Note. Convolution radius is 5 grid points, q25 and q75 are the first and third quartiles of the distribution. The rows next to the radar identifiers indicate the total number of objects in the warm (/) cold periods.

Almost all quantiles (excluding min) of the warm period are larger than the corresponding quantiles of the cold period, except for mean, q75, and max of the RATL radar and max of the RUDL radar marked in red. For the RAKU, RATL, RUDB, and RUDL radars, the q75 quartiles for the cold and warm periods are comparable.

We focus on the significant sizes of objects in observations and forecasts, for which we first discuss the general methodological and statistical problems of analyzing extreme values. In [Muraviev et al. 2022, 3] we present a more detailed and, if possible, formally accurate description of the extreme value theory (EVT) theorems, discuss the practical application of these theorems, and list some useful methodological and statistical recommendations.

2. Tail heaviness analysis: problems and some recommendations

With the rapid development of the extreme value theory and the abundance of its practical applications, most researchers are extremely cautious in their conclusions, avoid being categorical, and often emphasize the exploratory nature of the results obtained [Embrechts2003, Extremes and Integrated 2000, McNeil 2005, Novak 2011, Extreme events 2017]. Thus, [Paul Embrechts et al.2003] frankly warn the reader about possible failures calling some of their own graphs of shape parameter estimates "nightmarish". The most general recommendations are formulated, for example, in the fundamental work [Reiss, Thomas 2007] devoted to the statistical analysis of extreme values applied to insurance, finance, hydrology, and other areas.

Methods for estimating parameters are divided into *non-parametric* and *parametric* in the classification of Reiss and Thomas. In the first case, the object of analysis is a set of data on the basis of which sample distributions, densities and quantile functions are built. In the second case, the object of analysis is the generalized Pareto model, whose parameters (shape and scale) are estimated by different methods, including the methods of maximum likelihood, moments, Bayes methods, the Hill estimate. In the parametric approach, it is recommended to explicitly allocate the threshold value beforehand.

[Reiss, Thomas2007] consider the following steps for systematically solving the problem of statistical inference. First, consider the simplest *exponential model*, applying all the methods of parameters estimation mentioned above. If according to the diagnostic results (mostly graphical) it turns out to be unacceptable, then go to the *limited Pareto model* (with a zero threshold), supplementing the above parametric estimates with the Hill estimate. It is recommended to compare the Bayesian estimate (with the prior gamma distribution of the shape parameter) and the Hill estimate. In a good case, they should be close. For greater reliability, it is also recommended to conduct simulation modeling of random variables based on the same Pareto distribution. If diagnostics (for example, according to the graph of the sample mean function of kurtosis) shows a significant incorrectness of this model, then further considerations should be made before the next step. For example, the initial hypothesis about heavy tails, Pareto-type tails, could turn out to be incorrect, and then it makes no sense to move on to the full Pareto model. If the analysis did not

rule out "heavy-tailedness", then it is the turn of the full Pareto model with such methods for estimating parameters as the maximum likelihood, L-moments, probability-weighted moments, Bayes, etc.

A similar hierarchy has been proposed for the combination of the Pareto model with the Poisson distribution (Poisson-GP models).

The choice of thresholds is most reliably based on plots of *mean kurtosis values*, as well as on the analysis of the behavior of estimates of distribution parameters and their confidence intervals. In addition to a series of thresholds from some reasonable interval, one can experiment with "random thresholds" for order statistics at the end of the original variational series (e.g. [Galambos 1984]).

[Reiss, Thomas2007] recommend using the "pragmatic approach": we cannot know whether the resulting analysis of extremes will be useful for extrapolation outside the area in which the previous data is collected; "cross your fingers for luck" and publish a risk assessment of future extremes under the estimated distribution; update the model as more information becomes available.

3. Estimating Pareto Distribution Parameters Using the ExtRemes Package

We further reduce the term *generalized Pareto distribution* to *Pareto distribution*. Our approach is straightforward: the tail analysis method is parametric [Reiss]; distribution parameters are estimated by several methods; the location parameter (that is, the *Pareto threshold*) is fixed (the threshold on the precipitation object areas; the selection of informative thresholds is based on the analysis of histograms discussed below). Nonparametric methods for estimating heavy tails (for example, Hill estimators) are not used because of possibly large errors in small values of the shape parameter.

3.1. Pareto Distribution Parameter Estimates and Informative Threshold Selection

The approximation of the set of maximum object sizes by the Pareto distribution was performed using the `fevd()` function of the `extRemes` library from the R language repository. The main properties of the `extRemes` package including its theoretical and statistical principles are described in [Gilleland and Katz 2016].

In the function `fevd` for parameter estimates, the methods of 1) maximum likelihood, 2) generalized maximum likelihood, 3) L-moments and 4) Bayesian estimation based on statistical modeling of Markov chains are implemented. It is possible to include additional variables (covariates), for example, time, to take into account the seasonal cycle or to adjust the joint Pareto-Poisson model [Martins 2001, Gilleland, Katz 2011]. The choice of warm and cold periods for

testing made it possible, at least at this stage, to disregard the time covariate. Experiments with the joint Pareto-Poisson model led in some cases to strong computational instability, and it was decided to confine ourselves at this stage to the Pareto model. In [Muravyov 2022, 2], we give preliminary estimates for the distribution of peaks on the time scale to satisfy the conditions for the Poisson distribution.

3.1.1. Generalized maximum likelihood

The maximum likelihood method (MLM) is well known, so let us focus on its "generalization", as it was called by the authors of [Martins2001]; we present the basic formulas, using the specified source. If the Poisson distribution of time points with exceeding the threshold μ has intensity λ , then in formulas (1) the parameters are modified as follows:

$$\mu^* = \mu - \sigma(1 - \lambda^{\xi}) / \xi, \sigma^* = \sigma \lambda^{\xi} \text{ at } \xi \neq 0,$$

$$\mu^* = \mu + \sigma \ln(\lambda), \sigma^* = \sigma \text{ at } \xi = 0.$$

The quantiles of the Pareto-Poisson model are functions of the parameters and the return period (the expected time between exceedances):

$$x_p = \mu - \sigma[1 - (\lambda T_p)^{\xi}] / \xi \text{ at } \xi \neq 0,$$

$$x_p = \mu + \sigma \ln(\lambda T_p) \text{ at } \xi = 0.$$

The estimate of the intensity parameter is $\hat{\lambda}$ taken from the data and is equal to the exceedance frequency in the selected time interval, from which the original peak values are extracted. The log-likelihood function for the parameters θ of the Pareto-Poisson model with peaks $\{x_1, x_2, \dots, x_m\}$ exceeding the threshold x_0 from a set of n time units is written as

$$\begin{aligned} \ln[L(\theta | x)] &= \ln[P(M = m | \nu = \lambda n)] + \sum_{i=1}^m \ln[f(x_i | x \geq x_0)] = \\ &= m \ln(\lambda) - \lambda n - m \ln(\sigma) + \left(1 - \frac{1}{\xi}\right) \sum_{i=1}^m \ln(y_i), \quad (2) \end{aligned}$$

where $\theta = (\lambda, \sigma, \xi)$, the expression before the summation symbol represents the likelihood of observing exactly m values exceeding the threshold x_0 in an archive of length n time units, and the second part represents the likelihood that these m exceedances had observed values x_i ; $y_i = 1 + \xi(x_i - x_0)/\sigma$. The intensity λ of the Poisson process can be fixed, for example, by the sample frequency of exceeding the threshold, and thereby reduce the number of model parameters to two.

As is known, *the maximum likelihood method* applied to the scale and shape parameters leads to a system of two equations in two unknown variables, the solution of which is found using

computational algorithms of computational mathematics (in the ExtRemes package - by the Newton-Raphson and gradient methods). The *generalized maximum likelihood method* was proposed due to the fact that the standard MLM generated unnatural estimates of the shape parameter on small samples in the hydrological analysis of extrema. The authors [Martins 2000, 2001] used a "truncated" Bayesian approach, assuming that the shape parameter can be limited to reasonable limits (for example, by the usual interval for hydrology [-0.5, +0.5]) and made it a random variable with a beta distribution $B(p, q)$, $p=6$, $q=9$. The a priori distribution density of the shape parameter has the form $\pi(\xi) = (0.5 + \xi)^{p-1} (0.5 - \xi)^{q-1} / B(p, q)$, while the mean value is shifted to the positive region and is equal to 0.1, the variance is $(0.122)^2$. The generalized likelihood function was expressed as follows: $GL(\lambda, \sigma, \xi | \mathbf{x}) = L(\lambda, \sigma, \xi | \mathbf{x}) \pi(\xi)$. The approach is considered truncated, since the joint prior distribution is not used here. In formula (2) for the Pareto-Poisson model, the function $\pi(\xi)$ is added additively and the solution is found for the maximum value of the logarithm of the generalized likelihood function by the same methods of computational mathematics.

3.1.2. L-moments

To estimate the parameters σ and ξ the Pareto-Poisson model using *L-moments*, the main coefficients b_r are first determined using the rank series of the initial peak values (maximum areas of objects) $x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}$, where n is the number of elements in the sample (in our case, selected *situations* during the test period):

$$b_r = \sum_{i=1}^n \left[\frac{(n-1)(n-i-1) \dots (n-i-r+1)}{n(n-1)(n-2) \dots (n-r)} x_{(i)} \right], r = 0, 1, 2, \dots$$

The first three L -moments are calculated by the formulas:

$$l_1 = b_0, l_2 = 2b_1 - b_0, l_3 = 6b_2 - 6b_1 + b_0.$$

Using the notation $t_2 = l_2 / l_1$, we write the calculation formulas for the estimates:

$$\hat{\sigma} = l_1 \left(\frac{1}{t_2} - 1 \right), \quad \hat{\xi} = 2 - \frac{1}{t_2}$$

standard **method of moments**, these estimates are calculated using simple formulas:

$$\hat{\sigma} = \frac{1}{2} a \left(\frac{a^2}{s^2} + 1 \right), \quad \hat{\xi} = \frac{1}{2} \left(1 - \frac{a^2}{s^2} \right)$$

where a and s^2 are the sample mean and sample variance, respectively.

Formally, L -moments are defined using the formula

$$l_r = \frac{1}{r} \sum_{k=0}^{r-1} (-1)^k \binom{r-1}{k} \mathbf{E} X_{r-k:r}$$

where $X_{k:n}$ is the k - th smallest value in a sample of size n from the distribution of a random variable X , E is the expectation operator. In statistics, the first four L -moments are called L -mean (or L -position), L -scale , L -skewness , and L -kurtosis .

Convenient property of L -moments is the use of order statistics based on the values of a ranked series of realizations of the initial random variable, which, firstly, provides more robust statistical characteristics compared to the usual method of moments, and, secondly, guarantees the presence of higher moments at the only condition for the boundedness of the mathematical expectation. A consequence of this favorable feature is a lack of sensitivity: for example, the Laplace distribution has a kurtosis of six and light exponential tails, while the Student's distribution with three degrees of freedom has an infinite kurtosis and heavy tails; at the same time, the L -skewness for the Laplace distribution is higher than the same estimate for the indicated Student's distribution.

3.1.3. Bayesian method

This method is implemented in the ExtRemes package as a Markov chain simulation using Monte Carlo methods (Markov Chain Monte Carlo methods, MCMC). An algorithm for the automated generation of Markov chain states for symmetric *target* distributions was published in 1952 by a group of researchers at the Los Alamos Laboratory led by N. Metropolis, who conducted statistical computer experiments in the field of nuclear and thermonuclear weapons. In 1970, an article by W. Hastings appeared, in which the restriction by symmetric distributions was effectively removed, and at present this algorithm is called Metropolis-Hastings and consists of the following. Samples are taken from the probability distribution $P(x)$, the parameters of which are to be determined, and for which only the function $f(x)$ is known, which is proportional to the *target* probability function $P(x)$. The algorithm collects the states of the chain, which together approximate the desired distribution $P(x)$ with increasing accuracy. The selections are made iteratively with the only condition that each next value ("sentence") depends only on the previous value (hence the Markov property of the chain). The selection of the considered candidate for the completion of the sample depends on the comparison of the values of the function $f(x)$ taking into account the desired distribution $P(x)$.

The sequence of specific actions that reflect the essence of the *Bayesian strategy* is that a non-zero likelihood function $h(x)$ is used as a function $\theta f(x)$ and c , where θ is the vector of estimated parameters, and x is the initial sample of observations of length n . In its final form, $f(x)$ is the negative sum of n logarithms of the corresponding distribution densities of implementations of x under the condition θ .

The candidate selection function (*sampler*) is an unconditional and a priori specified distribution density of the parameter vector $g(\theta)$, acting as a matrix $g(\theta^*, \theta)$ of transition probabilities θ to θ^* .

fevd() calculation module implements the following iterative procedure.

1. The initial value θ^0 is fixed, usually close to some average position of the a priori distribution g , which is assumed to be normal by default (with parameters previously estimated by the maximum likelihood method). Using a random number generator tuned to the distribution g , a candidate θ^* is extracted, considered as the θ element of a *random walk following* θ^0 (the simplest version of the *Markov chain* [Feller1984, v.1]).

2. The *Hastings ratio* is calculated

$$r(\theta^0, \theta^*) = [h(x | \theta^*)g(\theta^*, \theta^0)] / [h(x | \theta^0)g(\theta^0, \theta^*)],$$

which eliminates the need to evaluate the unknown denominator (the integral of the likelihood over the probability measure of the parameter) in the Bayes formula for conditional probability.

3. An operation is performed, which is called the *failure of the Metropolis* :

+ calculated $a(\theta^0, \theta^*) = \min(1, r(\theta^0, \theta^*))$, treated as a *probability a*,

+ using the encoder of *uniformly* distributed numbers on $[0, 1]$, the number u is extracted and the decision is made: for $u \leq a$ the transition of the state θ^0 to θ^* is allowed, for $u > a$ the transition is not allowed and the new state coincides with the previous one.

Recursion at steps 2-3 leads to the generation of a sample of values that reproduce the desired distribution of the parameter $P(\theta)$, the *average* value (or *mode*) of which can be taken as a *Bayesian estimate of the* distribution parameter of the original random variable X . From the resulting sample of states of the Markov chain, *rank* estimates of the *reliable interval* (credible interval), which differs from the *confidence interval* (confidence interval) by the fact that in the first case the boundaries of the interval are fixed, and the parameter is variable, in the second case the boundaries are variable, and the parameter is constant.

Remark. In the Bayesian strategy for estimating parameters, a theory of optimal solutions is constructed using the so-called *conjugate prior distributions*, which greatly facilitate the solution of emerging problems [DeGroot 1974]. If the a priori distribution function (or density) of the random parameter distribution Θ is g , and the conditional distribution function of the random variable X at $\Theta = \theta$ is $f(x | \theta)$, then the posterior distribution $g(\theta | x)$ of the parameter Θ at $X=x$ is proportional to $g(\theta) f(x | \theta)$ for every θ . Then we say that the family of distributions g is *conjugate* to the family of distributions f .

Example. Let $X_1, \dots, X_n$ be a resampling of the Bernoulli distribution with an unknown parameter θ . Suppose the prior distribution of a parameter is a beta distribution with parameters $p > 0, q > 0$. Then the posterior distribution θ for $X_i = x_i$ ($i = 1, \dots, n$) is the beta distribution with parameters $p + y, q + n - y, y = \sum x_i$.

The simplification of the solutions of statistical problems lies in the fact that the desired a posteriori function does not require any parametric estimation - it is only necessary to modify the parameters according to simple formulas. Correspondence tables of distributions are compiled according to the principle of conjugation in Bayesian problems.

However, it is known that the postulation of a family of prior distributions of a parameter is the most vulnerable side of the Bayesian methodology, and the application of this methodology in practice, including, in particular, the contingency condition, must be accompanied by additional checks. In addition to the general critical problems of the Bayesian methodology, there are two significant statistical disadvantages of the particular MCMC algorithm. First, the samples are correlated, which is why even a large set of samples will not quite correctly reflect the desired distribution $P(x)$. Secondly, with the inevitable convergence of the Markov chain to the desired distribution, the initial samples may belong to a different distribution, which requires setting a *rejection period* (burn-in period).

In recent decades, many variations of the MCMC algorithm have been developed, designed to eliminate the most serious shortcomings, however, the simple and stable Metropolis-Hastings algorithm turned out to be by the end of the last century "a universal tool both in Bayesian inference and in solving numerous problems outside the Bayesian community" [Geyer 2011].

3.2. Description of the mathematical package extRemes

A detailed description of the extRemes package and training materials can be found in [Gilleland and Katz 2016, Gilleland 2020 I, II]. This package (in R terms, the *library*) contains general functions for analyzing extreme values with the possibility of including additional variables (*covariate*) and *declustering intervals* according to the methodology [Ferro and Segers, 2003]. `fevd()` and `ci()` are used as the main operational functions. Parentheses indicate the presence of configuration options, the equal sign indicates the specified values of these options.

The following *parameter estimation methods are provided*: 1) maximum likelihood (MLE), 2) L-moments, 3) generalized maximum likelihood method (GMLE), 4) Bayesian strategy (Bayesian). *Statistical inference* (calculation of confidence and confidence intervals) is carried out using 1) normal distribution, 2) likelihood profile, 3) Bayesian estimates for the posterior function,

4) bootstrap. Variants of *two-dimensional analysis* and *dependency testing* are connected: 1) dependence plot on the tails of distributions, 2) estimation of the extremality index according to [Geyer 2011].

Let us make some remarks about the criteria used for modeling quality. The Akaike criterion is calculated by the `fevd()` function, while for the Chi-square criterion, histograms were made by `hist()` of the R language, which implements the Sturges algorithm, which may not be optimal for extrema analysis due to the equidistant graduation of histograms. However, the use of the kernel smoothing histogram algorithm built into `fevd()` is unacceptable due to the coverage of the entire domain of definition, including values both after and before the Pareto threshold. The Akaike criterion is not calculated in the `fevd()` module for the L-moments and Bayes methods due to assessments of the model quality for these methods by other criteria (BIC and DIC). Chi-square is a common criterion for all methods of model quality.

3.2.1. Setting initial values and basic options for the `fevd()` function

In the case of MLE / GMLE, it is important to have good initial estimates of the parameters, while the `fevd` function tries to find these approximations itself (by default). The initial list specifies the initial values of the estimates for running a numerical optimization procedure (MLE / GMLE) or for MCMC iterations (Bayesian). By default, L -moment estimates and estimates based on the moments of the Gumbel distribution will be calculated; those estimates for which the log-likelihood takes the smallest negative value are used below. For the Bayesian method, it is recommended to test several initial values to make sure that they do not affect the final result. But if the initial values are not suitable, the standard MLE method is used.

Since there is some "spin-up" in the Monte Carlo simulation of the Markov chain, the first few hundred realizations should be removed from the resulting sample using the *rejection (burn-in) option*. By default, the number of steps in the random walk process is taken equal to 10000 with `burn-in = 500` of the first realizations.

To take into account temporal and calendar characteristics, the options `period` , `basis` and `time` are used; `units` specifies the base period (default is “ years ”) and time units (default is “ days ”), respectively.

For GMLE and Bayesian methods, it is possible to order *the prior distribution function* using the `priorParams` list. As mentioned above, by default for GMLE , the *beta distribution* is used in the interval from -0.5 to 0.5 with parameters $p = 9$ and $q = 6$ (at which the mathematical expectation is shifted to the positive region and equals 0.1), and default Bayesian estimation uses normal

distribution functions with MLE parameter estimates of means with a standard deviation of 10 for all parameters.

Only in the Bayesian method, there is a symbolic variable `proposalFun` that orders the name of the function that generates *the proposal parameters* at each iteration of MCMC. By default, a random walk chain is used: at the current value of the parameter, a *candidate is proposed* in the form of an additive to the current value of a normally distributed random variable with given parameters.

3.2.4. Calculating confidence intervals using the `ci()` function

The `ci()` function calculates confidence (and Bayesian *confidence*) intervals (CI) from the `fevd()` output, using the selected parameter estimation method in `fevd()` and the interval calculation method specified in `ci()`.

For L moments, the only method available in the package is a parametric bootstrap with the number of iterations `R`. It is recommended to determine `R` through trial and error, say by starting with `R = 100` and gradually increasing (by one or two hundred) until the results stabilize. By default `R = 100`.

For MLE/GMLE, when setting the `method = "normal"` option, the normal distribution approximation is used. If `method = "boot"`, then the parametric bootstrap is applied.

To calculate the CI of parameter estimates by the Bayesian method, the extreme percentiles from the resulting MCMC sample are used (after removing the first burn.in values).

Finally, `ci (method="profliker")` specifies the search for bounds on *the likelihood function profile*. Calculating CIs in this way is often the best tool for estimating the shape parameter and return periods when the distribution is skewed and the normal distribution is unsuitable. The likelihood profile is calculated based on the likelihood maximization for each individual parameter within a certain range of its values (option `xrange`) with fixed values of the remaining model parameters.

Let us briefly describe the *parametric bootstrap algorithm*.

- (1) Generate a sample of length `n` from the input data of the model being approximated.
- (2) Fit the distribution of the extreme value to this sample and store the resulting parameter estimates (and, for example, return periods).
- (3) Repeat steps (1) and (2) `R` times.
- (4) Based on the resulting sample of the previous steps, calculate confidence intervals using the corresponding extreme percentiles given by the `alpha` option.

3.2.3. Digital and graphic output

The *parameter estimates* and their *standard errors* are placed in a vector (`par`, `se.theta`), respectively. The *cross-covariance matrix* is contained in an array called `cov.theta`. Approximation *quality scores* are available under the names AIC (Akaike Information Criterion [Akaike, 1974]), BIC (Bayesian Information Criterion, [Schwarz, 1978]), and DIC (Deviation Information Criterion). The characteristics of the output information are specified by the options of the `summary` and `print` functions. In Bayesian Estimation, the type = “ trace ” option creates a plot panel indicating the posterior distribution function and a *trace panel of the MCMC method* for each parameter. By default, the resulting mean is calculated from the posterior sample as the parameter estimate, but setting the `FUN` = “ postmode ” option will plot the posterior sample *mode* .

3.3. Estimates of the quality of nowcasting using Pareto distribution parameters

The `fevd()` function, together with estimates of the scale and shape parameters, returns estimates of the standard error, which will be used below to calculate the boundaries of confidence intervals (L, U), the main tool for the precipitation nowcasting quality assessment in this study.

3.3.1. Archive of results of verification of nowcasting of precipitation areas

For each of the eight radars, the two periods of the year, and each of the Pareto thresholds, two tables of numerical data (sizes of maximum objects and histograms) and one set of graphs are built. For eight radars, two periods of the year, three Pareto thresholds, the total set is $8 \times 2 \times 3 = 48$ tables of object characteristics, 48 tables with histograms and 48 graphics files.

Output tabular data. Each table with area characteristics consists of six columns: 1) names, 2) observations on the locator, 3 - 6) forecasts according to the STEPS model - for 30, 60, 90 and 120 minutes. Two groups of characteristics are organized by rows: 1) general characteristics of the areas of objects and for four methods of estimating parameters 2) criteria for the quality of modeling (Akaike and Chi-square), and estimates of the scale and shape parameters with their confidence intervals.

Table 3 presents the general statistical properties *of object areas* according to the RUDL locator (Smolensk), with an area size of at least 625 points, for forecasts for four periods.

Table 3 . General characteristics of the areas of maximum objects in continuous precipitation situations according to observations (obs) of DMRL-S radar of Smolensk (RUDL) in the warm season (May - September 2017) and in STEPS nowcasting precipitation fields (30-120 min)

	obs	30	60	90	120
--	-----	----	----	----	-----

peaks	92	119	123	131	130
area_min	625	654	630	630	639
area_med	2496	2394	2380	2195	2150
area_max	19812	19982	20023	19579	18426
ndegf	11	11	12	11	11

Note. Row names: peaks - the number of maximum objects; area _ min / med / max - minimum, median and maximum area, respectively; ndegf - the number of bins on the corresponding histograms (the number of degrees of freedom for the Chi-square test, Sturges' rule).

The characteristics of Table 3 reflect the general properties of objects with a large area in the coverage areas of all radars. On the one hand, the number of objects with the maximum area in the prognostic fields exceeds the similar number in the observation fields by approximately 20-25%. On the other hand, in such statistics as the mean, median, and maximum values of these areas, the sets of observation fields and forecast fields are quite comparable.

The median area in the observation fields is 2496 points, and in the prognostic fields this area gradually decreases to 2150 points with the lead time; a similar trend is seen in the maximum values (decrease from 19812 to 18426). In this, one can see a trend in the distribution density of prognostic areas towards positive asymmetry, that is, more objects with sizes not exceeding the median and maximum values of the areas in the observation fields compared to the forecast. Such phenomenon reflects the main characteristic of most statistical models devoid of sources and sinks of energy: smoothing and corresponding suppression of extrema both in the values of quantities and in their areal characteristics, along with the splitting of connected objects into different parts.

Checking the statistical significance of the estimated scale and shape parameters should be preceded by an assessment of the quality of the simulation: if the quality of the model is unsatisfactory, no estimates of the model parameters can be taken seriously. However, some model quality metrics depend (explicitly or implicitly) on the model parameters, so that they also depend on the methods for estimating these parameters. For example, the Akaike information criterion is calculated using the likelihood function, which is used in the numerical methods of estimating MLE and GMLE, and therefore is a direct by-product of parameter estimates and depends on the features of the calculated cross-covariance matrix of parameters. In other methods, the covariance matrix of parameters is not used, and the quality of the simulation is assessed by other measures.

However, in the output of all methods, parameter estimates are available, which can be used to check their compliance with the original observations. The simplest, although not the most reliable way is to build histograms and calculate the Chi-square test. Consider, for example, the Akaike test and Chi-square as applied to samples of maximum areas, the characteristics of which

are shown in Table 3. Table 4 shows estimates of the scale and shape parameters of the Pareto distribution of the maximum areas of continuous precipitation observed during the warm period of 2017 in the fields of the Smolensk radar and in the fields of nowcasting. Parameter estimates are provided with confidence intervals (with a confidence level of 95%), which are calculated 1) from the standard error se for the MLE and GMLE methods by adding $\pm 1.62 se$ to the parameter estimate, 2) from the variation series of hundreds of bootstrap implementations for the L-moments method, and 3) from extreme percentiles of the resulting sample of the Markov chain for the Bayesian method .

The values of the scale parameter are given for information, since it is difficult to comment on them (on the Pareto density plot, the reciprocal of the scale is the value of the distribution density at conditional zero, in this case at point 625) and we will henceforth focus on the shape parameter. It can be seen that the distribution has a heavy tail - all shape parameter estimates are positive along with the confidence intervals (with a minor exception in the obs observation column where the lower CI is -0.043). The quality of the GPD (Generalized Pareto distribution) according to the Akaike criterion and according to Chi-square decreases with the transition to the forecast fields of 30 min. In both series of criteria values, there are "local" maxima: weak one at 90 min (Akaike) and noticeable one at 60 min (Chi-square) (in red in Table 4). The position and the value of the Chi-square maximum are most likely due to sample effects, and are associated, first of all, with the histogram construction technique and partially stochastic nature of the GMLE method. Such rare cases require separate consideration; for illustration, the values of the respective histograms are given in Table 5.

Table 4 . Quality criteria for modeling by the general Pareto distribution of the maximum areas of objects (Akaike and Chi-square), and estimates of the scale and shape parameters with the boundaries of their confidence intervals according to data similar to those in Table 3. The parameter estimation method is the generalized maximum likelihood (GMLE).

	obs	30	60	90	120
GML_Akaik	1661	2150	2225	2364	2352
GML_XI2	9.001	11.797	28.153	16.109	16.163
GML_scl_1	1471	1518	1531	1475	1564
GML_scl_2	2468	2270	2261	2155	2279
GML_scl_3	3464	3023	2991	2834	2995
GML_shp_1	-0.043	0.025	0.049	0.080	0.053
GML_shp_2	0.336	0.326	0.337	0.356	0.333
GML_shp_3	0.715	0.627	0.625	0.632	0.613

Note. Generalized Pareto distribution parameter estimation method - GMLE , Akaike - Akaike criterion, XI 2 - Chi-square criterion, scl - scale parameter estimation, shp - shape parameter estimation. Indices for parameter estimates: 1 - lower (2.5%) limit of CI, 2 - parameter estimate, 3 - upper limit (97.5%) of CI. Local maxima in the rows of quality criteria are highlighted in red.

Despite the fact that in the GMLE method the shape parameter has a prior beta distribution, which limits the values of the parameter to the interval $[-0.5, +0.5]$, the posterior distribution of the parameter does not have such restrictions, and the resulting estimate of the shape parameter will contain values outside the prior interval. In Table 4, the upper bounds of the confidence intervals noticeably exceed 0.5, which turns out to be quite a common feature of other radars as well.

Table 5 . The number of cases in gradations of the histogram of the distribution of maximum areas according to data similar to Table 4

observations/ forecasts (min)	histogram gradations										
	1	2	3	4	5	6	7	8	9	10	11
DMRL-S RU DL	38	23	10	7	5	3	3	2	0	1	
STEPS-030	55	29	10	9	7	3	2	3	0	1	
STEPS-060	57	30	10	10	4	6	0	5	0	0	1
STEPS-090	63	31	10	9	5	7	1	3	1	1	
STEPS-120	63	30	9	10	6	6	0	3	2	1	

Note. Initial data: DMRL-S Smolensk, warm period, Pareto threshold equals 625. Equidistant gradations were selected using the Sturges method.

Let us summarize the shape parameter data with the values of the Chi-square criterion in a separate table (Table 6). The degrees of freedom estimated by the number of histogram gradations and used for determining the critical values of the Chi-square test are between 6 and 12 for both periods of the year. The 5% critical values of the Chi-square test for the degrees of freedom in brackets are 12.592 (6), 16.919 (9), 18.307 (10), 19.675 (11), and 21.026(12).

The values of the Chi-square criterion in the forecast fields of 60 min have the character of outliers, while in the remaining rows of the table there is a quite regular change in the estimates of the shape parameter and their confidence intervals for all forecast periods without exception. Let us pay attention to the ranks of the parameter estimation methods. Peak objects in the fields of observations and forecasts at all times are modeled best by the Bayes method according to the quality criterion, worst by the L -moments method, while the shape parameter estimates turn out to be the largest mainly by the Bayes method and the smallest, by the L -moments method. All estimates of the shape parameter and confidence intervals, except for the lower bounds of the CI for the observation fields, are positive, which unambiguously confirms the “heavy-tailed” Pareto distribution of those areas of continuous precipitation whose area exceeds 625 grid points (which is equivalent to a square of 50 km × 50 km).

Table 6. Chi-square test and shape parameter estimates with confidence interval boundaries based on data similar to those in Table 4

	obs	30	60	90	120
GML_XI2	9.001	11.797	28.153	16.109	16.163
MLE_XI2	10.21	13.354	30.346	16.923	17.769
Lmo_XI2	10.777	14.121	32.105	18.581	19.092
Bay_XI2	7.773	11.125	26.734	15.132	15.404
best	<i>Bay_XI2</i>	<i>Bay_XI2</i>	<i>Bay_XI2</i>	<i>Bay_XI2</i>	<i>Bay_XI2</i>
worst	<i>Lmo_XI2</i>	<i>Lmo_XI2</i>	<i>Lmo_XI2</i>	<i>Lmo_XI2</i>	<i>Lmo_XI2</i>
GML_shp_1	-0.043	0.025	0.049	0.08	0.053
MLE_shp_1	-0.044	0.016	0.041	0.075	0.045
Lmo_shp_1	-0.007	0.034	0.047	0.069	0.052
Bay_shp_1	-0.008	0.037	0.063	0.11	0.091
larger	<i>Bay/Lmo</i>	<i>Bay/Lmo</i>	<i>Bay/GML</i>	<i>Bay</i>	<i>Bay</i>
smaller	<i>MLE/GML</i>	<i>MLE</i>	<i>MLE</i>	<i>LMO/MLE</i>	<i>MLE</i>
GML_shp_2	0.336	0.326	0.337	0.356	0.333
MLE_shp_2	0.294	0.289	0.309	0.34	0.304
Lmo_shp_2	0.278	0.274	0.29	0.312	0.284
Bay_shp_2	0.343	0.328	0.344	0.396	0.354
larger	<i>Bay/GML</i>	<i>Bay/GML</i>	<i>Bay/GML</i>	<i>Bay</i>	<i>Bay</i>
smaller	<i>Lmo</i>	<i>Lmo</i>	<i>Lmo/MLE</i>	<i>Lmo</i>	<i>Lmo/MLE</i>
GML_shp_3	0.715	0.627	0.625	0.632	0.613
MLE_shp_3	0.632	0.561	0.577	0.604	0.562
Lmo_shp_3	0.506	0.496	0.492	0.503	0.47
Bay_shp_3	0.831	0.681	0.676	0.731	0.698
larger	<i>Bay</i>	<i>Bay</i>	<i>Bay</i>	<i>Bay</i>	<i>Bay</i>
smaller	<i>Lmo</i>	<i>Lmo</i>	<i>Lmo</i>	<i>Lmo</i>	<i>Lmo</i>

Note. Initial data: the warm period of 2017, the Pareto threshold is 625. All methods for estimating parameters were used - GMLE , MLE , L _ moment , Bayesian . Indices for parameter estimates: 1 - lower (2.5%) limit of CI, 2 - parameter estimate, 3 - upper limit (97.5%) of CI. The values of the Chi-square test that do not contradict the acceptance of the hypothesis about the validity of the generalized Pareto distribution model are highlighted in green; cases of excess of the values of the Chi-square test for the 95% significance level are highlighted in red (local maxima on the 60-minute forecast). Oblique font marks the ranks of methods according to the principle of the largest and smallest values of the corresponding characteristics.

Comparing the estimates of the Chi-square criterion, we can state that, in terms of the quality of modeling, all values of the criterion, except for forecasts for 60 minutes, do not reject the hypothesis of the applicability of the generalized Pareto distribution model to selected objects.

Ranking the methods by the Chi-square criterion provides the following information: the best quality of modeling for a given radar, a given period of the year and for a specified Pareto threshold is provided by the Bayesian method, the use of the L-moments method leads to the worst quality of modeling. For other radars, there is a wider variety of rankings, making it difficult to generalize preferences or systematic features to all radars and to all lead times.

Table 7 gives the shape parameter characteristics for all radars (Table 7). For brevity, we exclude confidence intervals and consider the warm period data with a Pareto threshold of 625 and using the Akaike criterion of the quality of the shape parameter estimate.

Table 7. Sample volumes of maximum precipitation areas (peaks) and evaluation of the quality of modeling of the generalized Pareto distribution (GPD) (625 grid points threshold) by the Akaike criterion for radars of the Central Federal District of Russia, observations (obs) and forecasting (30-120 min) in the warm period of 2017. Method parameter estimates - GMLE

DMRL-S		observations DMRL-S radar	forecast period, STEPS model (min)				
		obs	30	60	90	120	
Kursk, RAKU	peaks	86	117	118	121	123	
	Akaike	1496	2098	2123	2179	2219	
Tula RATL	peaks	79	117	117	128	132	
	Akaike	1463	2103	2115	2298	2367	
Voyekovo RAVO	peaks	89	133	141	145	157	
	Akaike	1620	2381	2534	2609	2816	
Bryansk RUDB	peaks	76	90	96	99	99	
	Akaike	1360	1609	1709	1760	1768	
Kursk RUDK	peaks	96	130	142	145	155	
	Akaike	1740	2346	2564	2631	2800	
Smolensk RUDL	peaks	92	119	123	131	130	
	Akaike	1661	2150	2225	2364	2352	
Nizhny Novgorod RUDN	peaks	84	122	129	141	147	
	Akaike	1713	2415	2379	2588	2689	
Valdai RUWJ	peaks	85	119	128	138	152	
	Akaike	1582	2202	2359	2533	2772	

As noted in [Muraviev2022], the number of objects, and, hence, the number of objects of significant and maximum area in the prognostic fields, gets overforecasted noticeably already at the first steps of nowcasting (up to 30 min), it also grows in other periods, although not so much. Let's consider, for example, the ratio of the number of objects at the forecast step of 120 min to the number of objects in observations ($123/86 = 1.43$, etc.) and compare it with the respective ratio of the values of the Akaike criterion ($2219/1496 = 1.48$, etc.). A slightly surprising fact is revealed: the larger the sample, the worse the quality of modeling this sample using the Pareto distribution. Thus, for the Bryansk (RUDB) radar, a 130% increase in the sample size corresponds to the same decrease in the modeling quality. The values of the Akaike criterion for the RUDB locator data turn out to be the best for all columns of the table.

Using a similar scheme, we composed a table of shape parameter values estimated by all four methods (Table 8).

Table 8 . Estimates of the shape parameter of the generalized Pareto distribution (GPD) (at a threshold of 625) for the Central Federal District radars based on observational data (obs) and forecasting (30-120 min) in the warm period of 2017

DMRL-S	GPD evaluation method	observations DMRL-S	forecast period, STEPS model (min)			
		obs	thirty	60	90	120
Kursk RAKU	GMLE	0.428	0.405	0.413	0.390	0.366
	MLE	0.472	0.419	0.434	0.394	0.353
	Lmom	0.379	0.354	0.360	0.336	0.312
	bayes	0.681	0.560	0.628	0.533	0.447
Tula RATL	GMLE	0.278	0.360	0.344	0.376	0.371
	MLE	0.161	0.342	0.316	0.370	0.363
	Lmom	0.181	0.307	0.290	0.329	0.324
	bayes	0.179	0.424	0.361	0.432	0.436
Voyekovo RAVO	GMLE	0.228	0.294	0.307	0.326	0.314
	MLE	0.074	0.241	0.264	0.294	0.280
	Lmom	0.094	0.238	0.255	0.275	0.267
	bayes	0.109	0.260	0.273	0.341	0.307
Bryansk RUDB	GMLE	0.327	0.303	0.328	0.348	0.322
	MLE	0.271	0.236	0.284	0.319	0.276
	Lmom	0.266	0.237	0.275	0.301	0.268
	bayes	0.342	0.280	0.355	0.397	0.311
Kostroma RUDK	GMLE	0.284	0.331	0.328	0.303	0.361
	MLE	0.199	0.299	0.298	0.259	0.350
	Lmom	0.206	0.281	0.281	0.252	0.316
	bayes	0.234	0.355	0.336	0.288	0.417
Smolensk RUDL	GMLE	0.336	0.326	0.337	0.356	0.333
	MLE	0.294	0.289	0.309	0.340	0.304
	Lmom	0.278	0.274	0.290	0.312	0.284
	bayes	0.357	0.322	0.354	0.385	0.336
Nizhny Novgorod RUDN	GMLE	0.000	0.000	0.253	0.308	0.322
	MLE	0.000	-0.002	0.165	0.262	0.287
	Lmom	0.038	0.153	0.180	0.250	0.267
	bayes	0.038	0.135	0.167	0.307	0.320
Valdai RUWJ	GMLE	0.287	0.249	0.262	0.291	0.327
	MLE	0.190	0.153	0.184	0.237	0.296
	Lmom	0.202	0.169	0.194	0.234	0.276
	bayes	0.203	0.162	0.196	0.247	0.318

Note. Pareto distribution parameters are estimated using generalized maximum likelihood (GMLE), maximum likelihood (MLE), L-moments (Lmom), and Markov chain simulations using the Metropolis-Hastings algorithm (Bayesian).

The main property of the shape parameter estimates – positivity and a noticeable difference from zero – indicates “heavy” tails of the distribution. The only exception is the data for the Nizhny Novgorod radar. It is the only case of almost zero and even slightly negative values of the shape parameter for observations and for forecasts for 30 min. However, the Akaike estimates in Table 7 on the RUDN radar line do not differ significantly from estimates for other radars.

When switching to a 30-min forecast, the value of the shape parameter can either increase (RATL, RAVO, RUDK) or decrease (RAKU, RUDB, RUDL, RUWJ), for all methods of parameter estimation.

Comparison of the shape parameter estimates obtained by different methods can be carried out by the minimum and maximum values in the corresponding columns. So, for example, in the data for the RAKU locator in the observation column, the maximum value of the shape parameter is 0.681 and obtained by the Bayes method, and the minimum value is 0.379 and obtained by the L - moment method. The ranking of methods in all columns of the table led to the following statistics. The highest shape parameter values are provided by the Bayesian methods (21 cases) and the generalized maximum likelihood method (19 cases); the minimum values of the shape parameter were obtained mainly by the L-moment method (30 cases) and only in 10 cases by the maximum likelihood method. The reasons for the similarity of the methods that generate the maximum shape parameter estimates are quite obvious: both are based on the Bayesian approach, either entirely (Bayesian), or in a "truncated" form (GMLE). As a rule, estimates by the MLE method are lower than those of the related GMLE method. Interestingly, there are cases of maximum discrepancy between the MLE and GMLE estimates: in the observational data of RATL (0.161 and 0.278), RAVO (0.074 and 0.228), RUDK (0.199 and 0.284), RUWJ (0.190 and 0.287), as well as in two or three columns predictive data.

Output graphic data. The graphical output reveals other aspects of modeling large areas of precipitation using the generalized Pareto distribution. The graphics for each radar and each Pareto threshold are displayed in one pdf file, consisting of 20 pages (slides) as follows. For each of the four methods, there are five pages corresponding to an observation and four predictions. Each of the pages for GMLE, MLE and L-moments methods contains four panels: 1) quantile diagram "model-observation", 2) quantile diagram with regression line "observation-simulation", 3) distribution density of observations (in kernel smoothing) and model Pareto density, 4) return period (in conditional "years") with confidence intervals. Each of the pages for the Bayesian method contains four panels: 1) and 2) the posterior distribution density of the scale and shape parameters, respectively (in kernel smoothing), 3) and 4) "traces" - implementations of the Markov sequence for the posterior distribution densities of the scale and shape parameters (10 000 steps, with the level of "burnout" in burn in = 500 steps).

For *modeling quality estimates* all the criteria, which are available in the fevd module (AIC , BIC and DIC) were partially analyzed, criterion χ^2 was additionally calculated for histograms built by the function hist () of the R language. Unfortunately, none of the built-in quality criteria apply to

all four parameter estimation methods. However, since the Akaike criterion, assuming a normal distribution of model errors, coincides up to an additive constant with the number $\chi^2 + 2k$, where k is the number of parameters in the model (in our case, always $k = 2$), then for a *comparative assessment* of the quality of modeling it is perfectly acceptable to confine ourselves to criterion χ^2 in its simplest form. Recall that the Akaike information criterion is calculated by the formula $4k - 2\ln(L)$, where L is the maximum of the likelihood function.

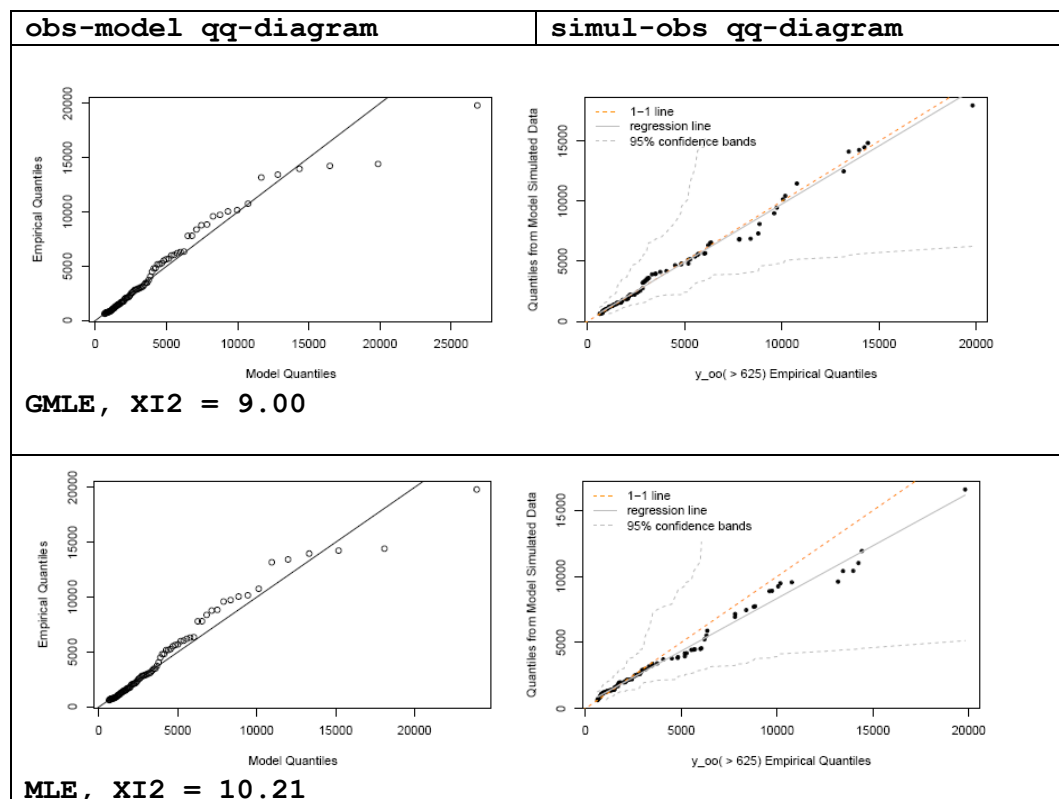
Let's illustrate the data of Tables 3-6 for the Smolensk locator (RUDL) with examples of the graphical output of the fevd() module, limiting ourselves to observations and forecasts for 90 minutes. Figures 2 and 4 show the quantile diagrams of the GMLE, MLE and L-moments methods, Figures 3 and 5 show the *posterior distribution densities* of the scale and shape parameters along with their *traces of the random walk* by the Bayesian method. Quantile diagrams are built in two ways: 1) between observations and the model data restored from the corresponding quantiles by the inverse transformation of the generalized Pareto distribution with estimated scale and shape parameters; 2) between observations and simulated data generated by a random number generator tuned to a Pareto distribution with specified parameter estimates. On Figures 2 and 4, the first set of diagrams (obs - model) is placed in the left column of panels, and the second set of diagrams (simul - obs) is placed in the right column of panels.

The first and, perhaps, the main remark should immediately be made about the *objects* of the largest sizes that were stored in the samples of *peaks above the threshold*, although sometimes such objects were precipitation areas that went beyond the survey circles of the locators, violating the physical, and hence the statistical correctness of the sample. The exclusion of such objects (*selection censoring*) is provided by the max option size in the FeatureFinder() function of the SpatialVx library used to generate the spatial objects for verification [Ants2022], however, it was difficult to do this in an automated mode. One can understand the need for censoring when analyzing the quantile plots in the left columns of Figs. 2 and 4. For example, a greater agreement between the model and observed data could be achieved by excluding objects of more than about 20000 points in size, which is equivalent to a square with a side of about the radar viewing circle radius (250 km).

Despite the presence of outliers in the quantile plots of the left column, it is possible to obtain, albeit with a wide range of confidence intervals, an acceptable match with the simulation data. For example, in Fig. 2, a noticeable deviation from the line of coincidence of one or two extrema in the diagrams of the left column is compensated by quite successful imitation in the diagrams of the right column for the GMLE and MLE methods. This, unfortunately, cannot be said about the data for the 90-minute forecast fields (Fig. 4), where the quality of modeling according to the Chi-square

criterion is noticeably inferior to the quality of modeling of observation objects (Fig. 2). It is obvious that when the sizes of objects are censored by the upper bound (about 15 thousand points), the diagrams of both columns will be more attractive.

The posterior distribution densities and traces of the random walk of the estimates of the scale and shape parameters are shown in Figs. 3 for objects in the observation fields and in Fig. 6 for objects in the forecast fields for 90 min (Fig. 6). On the trace graphs, the *burn-in line is highlighted*, which excludes from the construction of a posteriori densities all realizations of the Markov chain to the left of the point 500. Thus, the full set is 9500 realizations of the randomized parameters of the Pareto distribution. Comparing Fig.3 and 5, we note the following. Density modes for scale and shape in predictive samples deviate in different directions - the scale decreases by about 500 points (the Pareto distribution density curve at "zero" rises), and the shape increases from about 0.4 to 0.5 (the tail of the Pareto distribution density "gets heavier") . In particular, this behavior of the Pareto distribution density curve suggests that, when forecasting for 90 minutes, the STEPS model in the vicinity of the conditional "zero" (equivalent to the threshold of 625 points) systematically overestimates, then underestimates, and starting from a certain size of objects, begins to systematically overestimate the areas of objects. An increase in the shape parameter to 0.5 means, in particular, that the Pareto distribution for predicted objects can only have a mathematical expectation, and the variance estimate based on such data will lose its meaning.



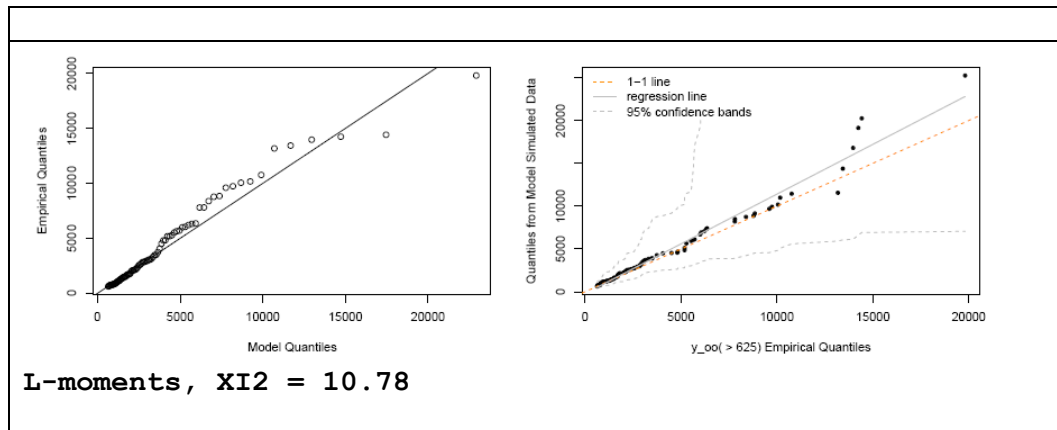


Fig.2. Quantile charts for comparing observational data and Pareto distribution simulation data with estimated scale and shape parameters (obs - model, left panel column), as well as simulation data with stochastic sampling on estimated Pareto distribution parameters (simul - obs, right panel column). Pareto distribution threshold - 625 points. Parameter estimation methods: generalized maximum likelihood (GMLE), maximum likelihood (GMLE), L - moments (L - moments). Radar data Smolensk (RUDL), warm period 2017

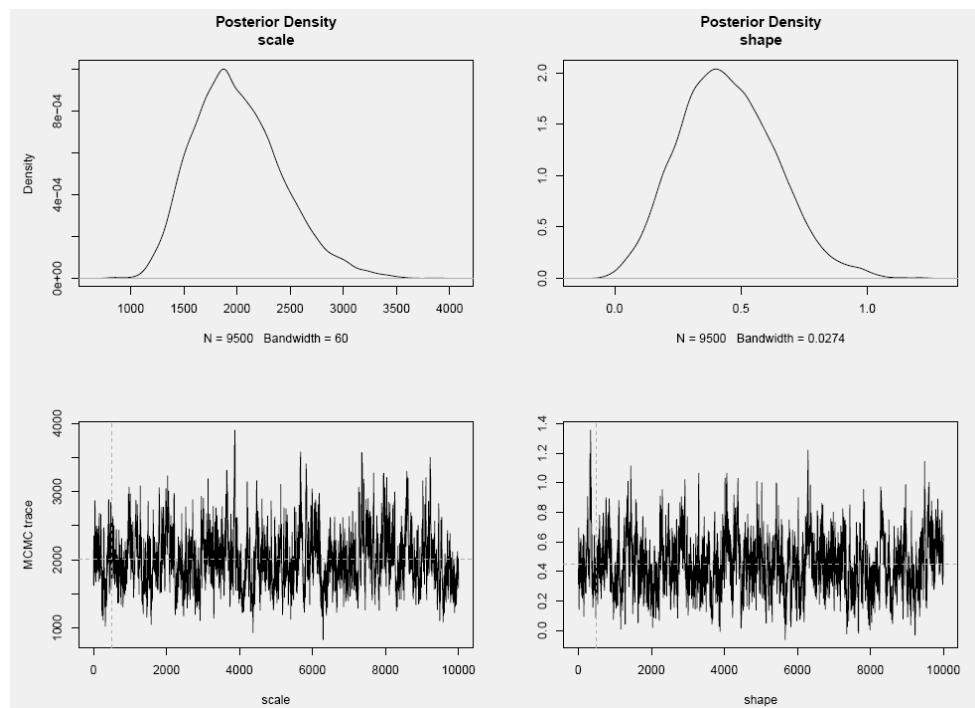


Fig.3 . Posterior distribution densities of the scale parameter (upper left panel) and shape parameter (upper right panel) and random walk traces of the scale parameter (lower left panel) and shape parameter (lower right panel) plotted by Bayesian method using Markov chain generation Metropolis-Hastings (Monte Carlo) algorithm. Chi-squared test = 8.3. Radar data Smolensk (RUDL), warm period 2017

obs-model qq-diagram	simul-obs qq-diagram

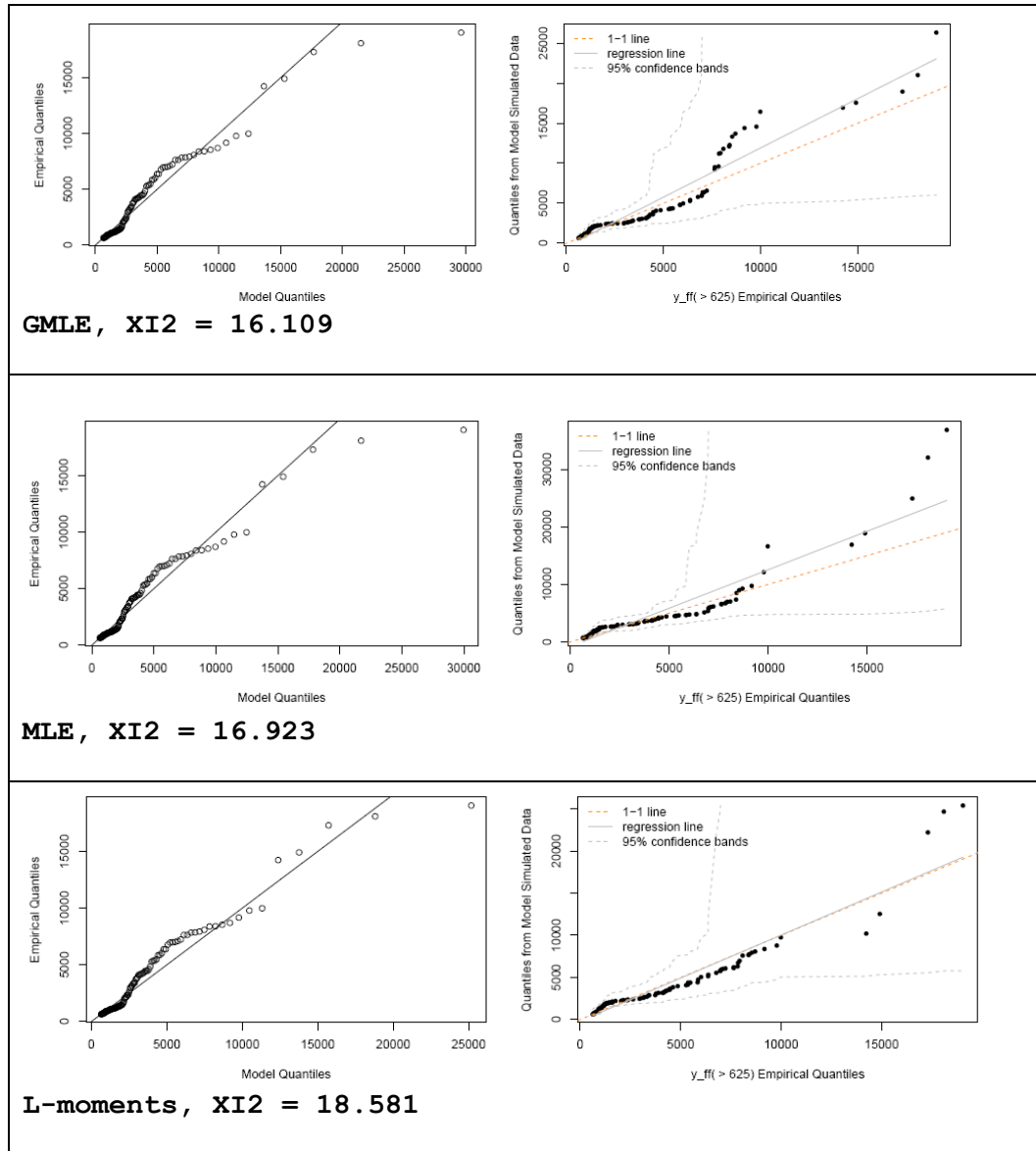


Fig.4. Quantile charts for comparing observational data and Pareto distribution simulation data with estimated scale and shape parameters (obs - model , left panel column), as well as simulation data with stochastic sampling on estimated Pareto distribution parameters (simul - obs, right panel column). Pareto distribution threshold - 625 points. Parameter estimation methods: generalized maximum likelihood (GMLE), maximum likelihood (GMLE), L-moments. STEPS model forecasts for 90 min, field of view and initial data of the Smolensk (RUDL) radar, warm period 2017

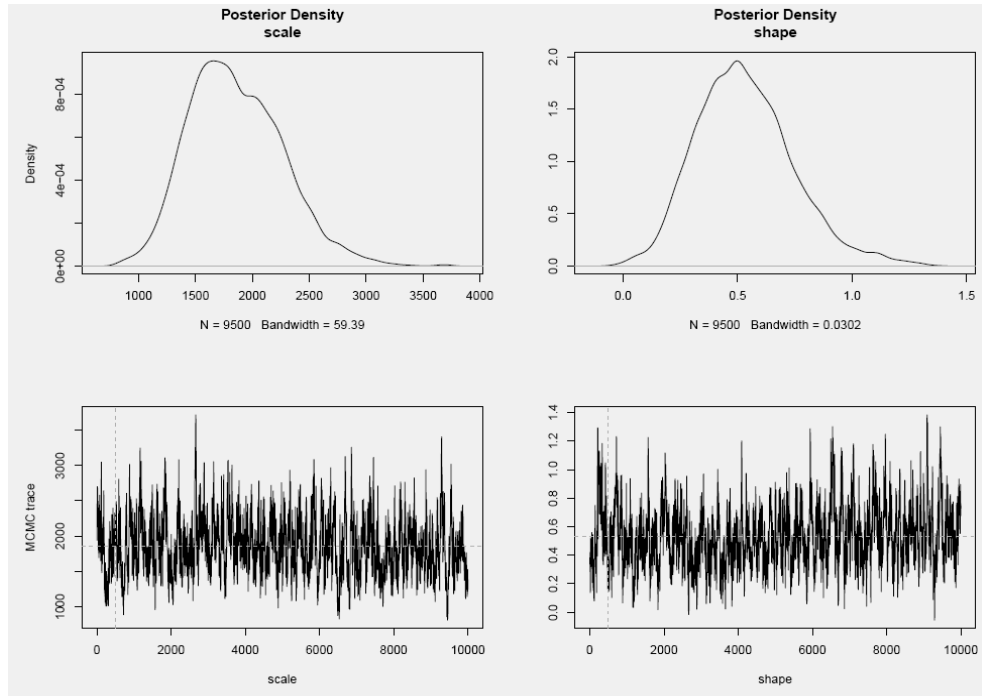


Fig.5. Posterior distribution densities of the scale parameter (upper left panel) and shape parameter (upper right panel) and random walk traces of the scale parameter (lower left panel) and shape parameter (lower right panel) plotted by Bayesian method using Markov chain generation Metropolis-Hastings (Monte Carlo) algorithm. Chi-squared test = 14.885. STEPS model forecasts for 90 min, field of view and initial data of the Smolensk (RUDL) radar, warm period 2017

3.3.2. Choice of parameter estimation method and Pareto distribution threshold

It is necessary to limit the number of methods for estimating distribution parameters and to reduce the set of thresholds to obtain verification results of nowcasting objects of a large area common for all radars. Based on the shape estimates obtained and their confidence intervals, it can be argued that all four methods led to consistent conclusions about the shape parameter for the threshold of 625 grid points. The standard method of estimating is undoubtedly the maximum likelihood method (MLE) associated with a modified Chi-square minimum method [[Kramer1975](#)], which, as mentioned above, can also replace the Akaike information criterion under some general assumptions. However, on small samples, MLE can lead to unnatural parameter estimates, which has led to the suggestion of a truncated Bayesian correction (GMLE) [[MartinsSted 2000, 2001](#)]. The L-moments methods are attractive due to the simplicity of calculations and the statistical robustness of the estimates. However, in [[MartinsSted 2001](#)], statistical experiments with the estimation of large quantiles showed the advantage of the GMLE method over the L-moments for samples of medium size and heavy tails, i.e. when the shape parameter is at least positive. Full confidence in the Bayesian parameter estimation strategy is hindered by a lack of experience in the broad sense, including insufficient mastery of the methodology, and experience in applying this strategy to extreme values in particular. The existence of many methods for estimating parameters

confirms, on the one hand, the complexity of the statistical analysis of extrema, and, on the other hand, excludes the existence of one general and universally applicable method. In this regard, we will use the GMLE method to bring the estimates of the nowcasting quality to a small number of observable results.

Let's try to solve the issue of thresholds choice for the Pareto distribution using histograms and Chi-square estimates, combining the graphic images with the values of the criterion assessment of the modeling quality.

Figure 6 shows the histograms of the size distribution of objects no smaller than 625, 900, 1225, and 1600 points and the GPD density values connected by linear segments. Parameter Estimation Method is GMLE. The titles of the panels indicate the sample sizes, number of gradations automatically calculated, extreme values and median sizes, as well as the parameter estimates. Differences and similarities between the histograms and approximating Pareto distribution density curves are visually visible for objects with sizes above the thresholds of 625, 900, 1225 and 1600 points. Let us recall that, on average, the number of objects in the forecast fields is bigger than in the observation fields. Since a larger Pareto threshold selects a subset of the maxima selected for a smaller threshold, the approximation of the higher-threshold subset by the Pareto distribution must increase the scale (going to the right along the tail) and change the shape.

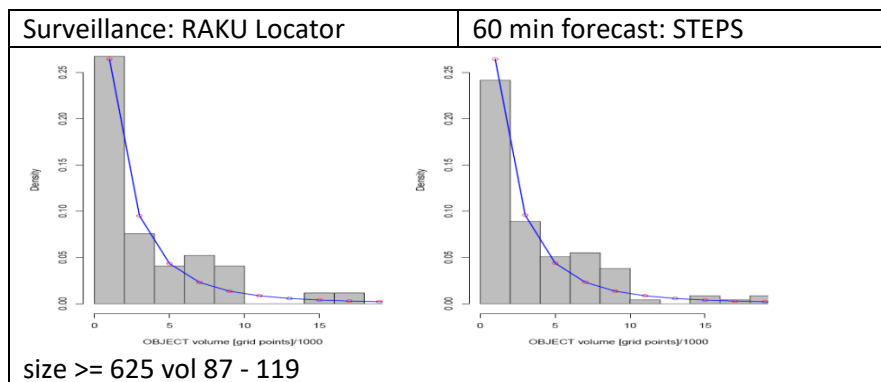
Let us write out the estimates of the χ -square criterion for Fig. 6 for the warm period:

$$625: \chi^2(\text{RAKU}) = 13.190, \chi^2(\text{STEPS -60}) = 14.721,$$

$$900: \chi^2(\text{RAKU}) = 13.106, \chi^2(\text{STEPS -60}) = 14.201,$$

$$1225: \chi^2(\text{RAKU}) = 13.420, \chi^2(\text{STEPS -60}) = 15.788,$$

$$1600: \chi^2(\text{RAKU}) = 14.099, \chi^2(\text{STEPS -60}) = 24.245.$$



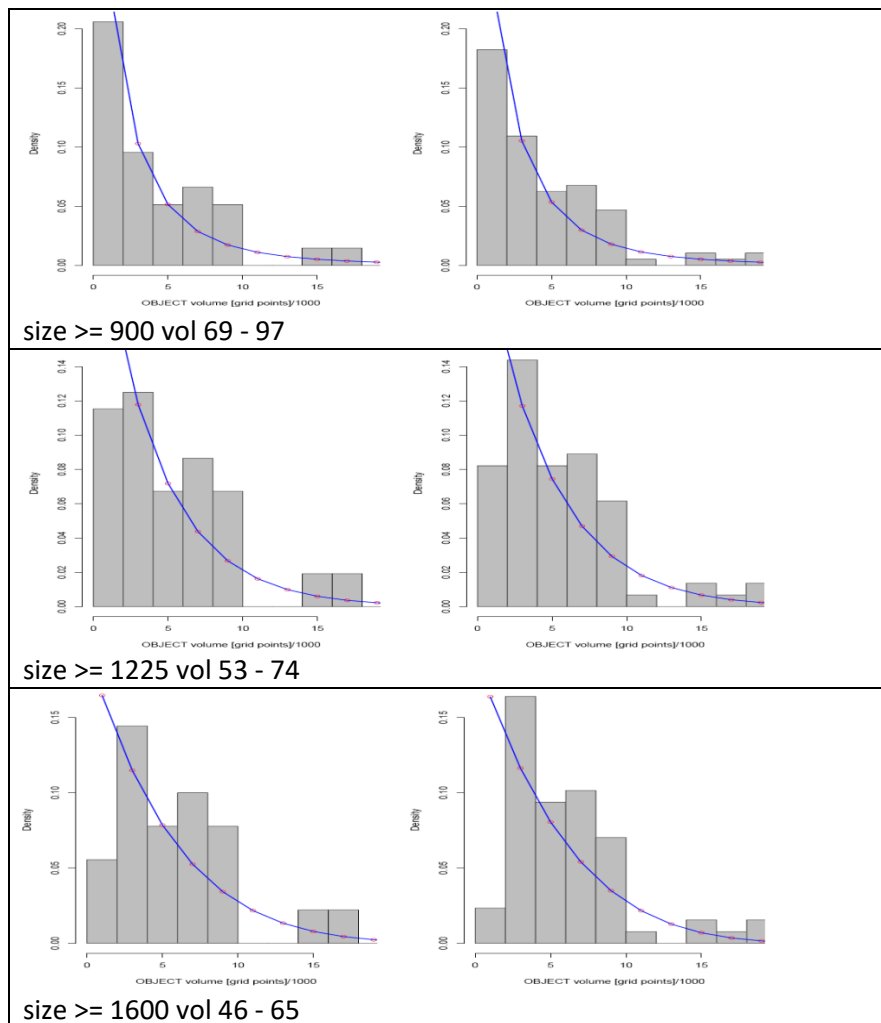
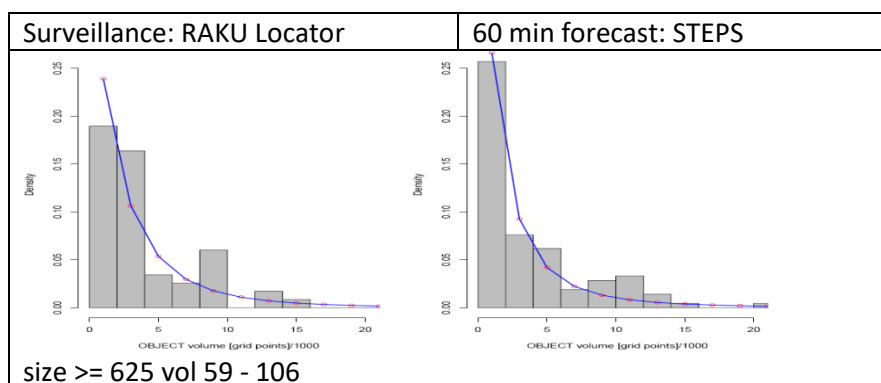


Figure 6. Warm period . Histograms and Pareto distribution approximation of object sizes in precipitation fields in Kursk radar observations (RAKU, left column) and forecasts (STEPS-60, right column) for 60 min. The Pareto threshold is (from top to bottom) 625, 900, 1225 and 1600 field points. Dimensions are given in size/1000 scale.



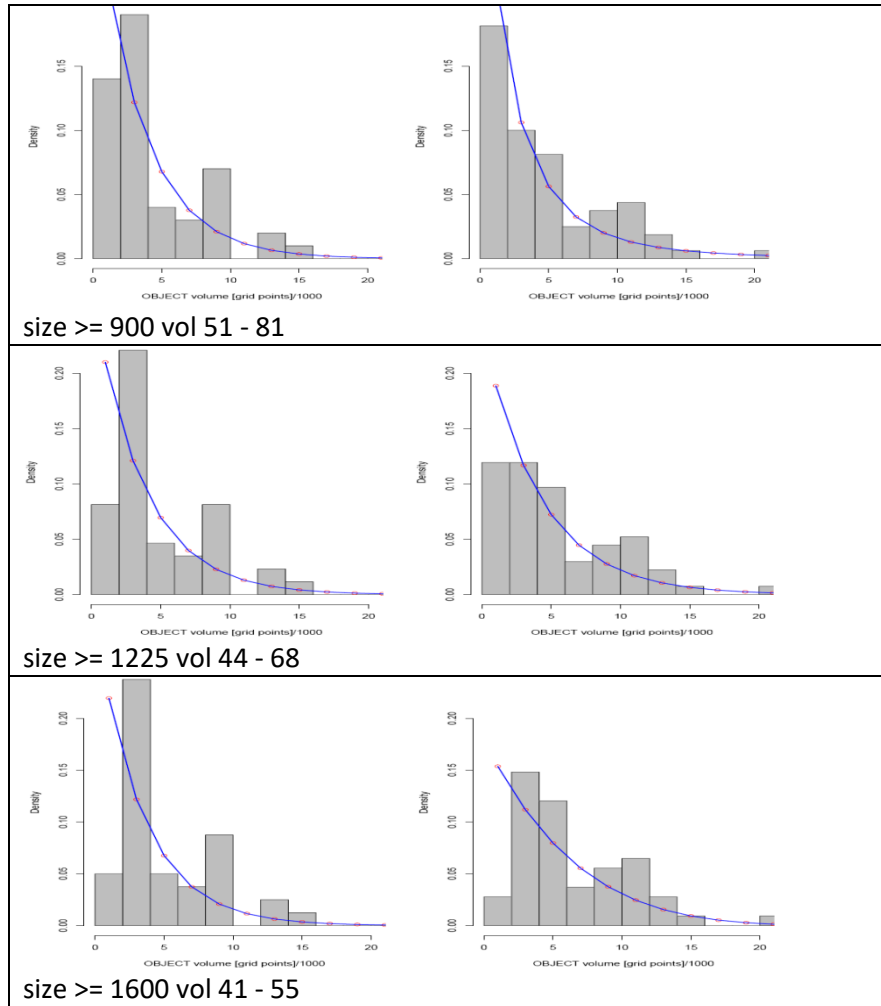


Figure 7. Cold period. Histograms and Pareto distribution approximation of object sizes in precipitation fields in Kursk radar observations (RAKU, left column) and forecasts (STEPS-60, right column) for 60 min. The Pareto threshold is (from top to bottom) 625, 900, 1225 and 1600 field points. Dimensions are given in size/1000 scale.

Let us write out the estimates of the χ -square criterion for Fig. 7 for the cold period:

$$625: \chi^2(\text{RAKU}) = 10.849, \chi^2(\text{STEPS -60}) = 13.646,$$

$$900: \chi^2(\text{RAKU}) = 12.427, \chi^2(\text{STEPS -60}) = 10.810,$$

$$1225: \chi^2(\text{RAKU}) = 16.501, \chi^2(\text{STEPS -60}) = 10.982,$$

$$1600: \chi^2(\text{RAKU}) = 22.069, \chi^2(\text{STEPS -60}) = 15.124.$$

For the Kursk radar for the period May-September 2017, the threshold size of an object of the about 600-900 grid points is the most suitable for modeling with generalized Pareto distribution, while objects with a size of 1200 and above are unsatisfactorily modeled with GPD. A similar analysis for the rest of the radars makes it possible to generalize this conclusion to all radars in the

Central Federal District of Russia, which means that using the GPD parameters for verification purposes is justified under the specified limitations.

Along this path, one can approach the solution of the issue of choosing an appropriate threshold (for the specified method of estimating the distribution parameters) for the resulting estimates of the quality of the nowcasting, that is, assessing the correctness of the construction of samples for applications of the Pareto distribution. This choice is based both on the visual representation (Figures 6 and 7) and on the values of the Chi-square test (Table 9)

Table 9 summarizes Chi-square estimates for object areas in radar fields (RADAR columns) and in 60 min forecast fields (STEPS -60 MIN columns) for tests in warm and cold periods of the year. In the hist() function, the number k of gradations is determined using *the Sturges rule* : $k = 1 + [\lg_2(n)]$, where n is the sample size. The analysis was carried out according to the joint data of observations and forecasts for each period and for each observation-forecast pair.

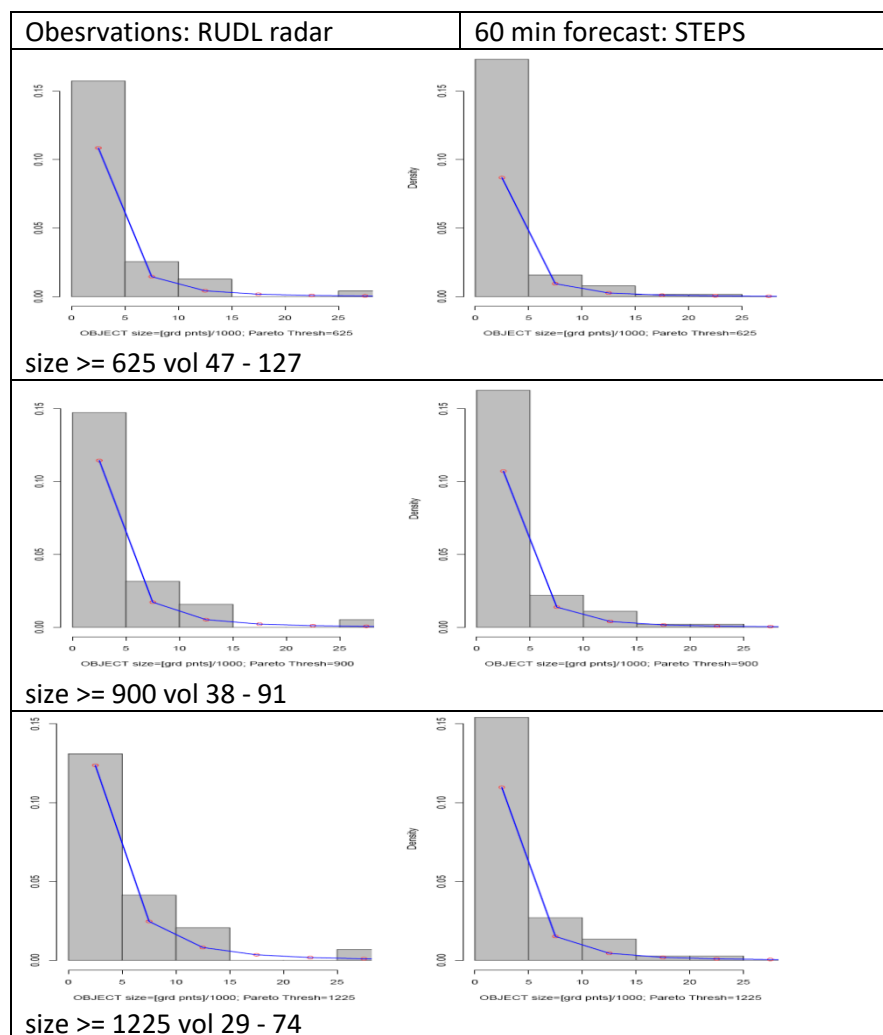
Table 9. Chi-square test values for assessing the quality of histogram approximation by the Pareto distribution with estimated scale and shape parameters at thresholds of 625, 900, 1225, and 1600 points.

	WARM PERIOD							
	RADAR				STEPS-60 MIN			
RADAR(cases; ndeg)	625	900	1225	1600	625	900	1225	1600
RAKU (87-46;11)	13.190	13.106	13.420	14.009	14.721	14.201	15.788	24.245
RATL (80-57;11)	4.097	5.674	8.708	12.476	11.395	12.140	10.565	17.086
RAVO (90-65;11)	6.921	8.417	13.148	19.733	4.637	7.423	11.288	21.318
RUDB (77-52;11)	6.690	9.289	13.859	21.479	7.695	10.024	15.256	21.446
RUDK (97-66;10)	6.862	8.432	13.818	18.803	5.947	8.044	13.332	20.265
RUDL (93-61;12)	4.711	6.310	9.959	13.958	14.335	16.637	22.859	30.741
RUDN (85-61;13)	4.301	5.301	9.213	12.080	7.474	10.088	14.666	22.267
RUWJ (86-61;11)	4.059	4.499	9.204	12.638	5.220	7.727	13.841	21.949
	COLD PERIOD							
	RADAR				STEPS-60 MIN			
RADAR	625	900	1225	1600	625	900	1225	1600
RAKU (59-41;12)	10.849	12.427	16.501	22.069	13.646	10.810	10.982	15.125
RATL (48-25;13)	10.775	10.747	11.624	11.623	13.599	12.759	10.450	7.272
RAVO (54-29;10)	10.296	11.282	15.997	23.012	8.947	9.096	15.831	23.653
RUDB (46-25;10)	7.617	8.678	11.632	9.365	7.689	7.192	10.107	14.690
RUDK (41-18;10)	11.502	13.442	13.641	16.047	12.697	12.139	15.109	22.009
RUDL (47-25;7)	3.604	3.020	2.056	1.888	12.928	4.424	3.651	2.909
RUDN (41-20;11)	12.630	11.117	10.904	12.481	6.423	5.255	7.355	10.534
RUWJ (27-16;12)	8.370	10.298	14.593	24.938	13.222	12.296	13.484	18.493

Note . In the RADAR column, next to the identifier, in brackets are indicated ("sample size of observations" for a threshold of 625 points - for a threshold of 1600 points; an estimate of the number of degrees of freedom).

The yellow background highlights the values that exclude the Pareto distribution from the set of suitable approximations; on the histogram, this is, as a rule, the second bin being larger than the first one, that is, the violation of the characteristic Pareto distribution density curve.

The values highlighted in red reflect, at first glance, one of the most important conditions of the second extremum theorem, *threshold stability*: the larger the threshold, the more accurately the data is modeled by the Pareto distribution. In real samples of a limited size rapidly decreasing with increasing threshold, this phenomenon should be recognized as a rare success. Let us consider in more detail the corresponding observational data and forecast for 60 min in the field of view of the Smolensk radar during the cold period. Figure 8 contains histograms of the sizes of the maximum objects in the fields of observations and forecasts for 60 min with a size of 625, 900, 125, and 1600 grid points. The decrease in the value of the Chi-square criterion with increasing threshold is ensured by an increasingly accurate approximation of the Pareto distribution density curve of the first bin of histograms.



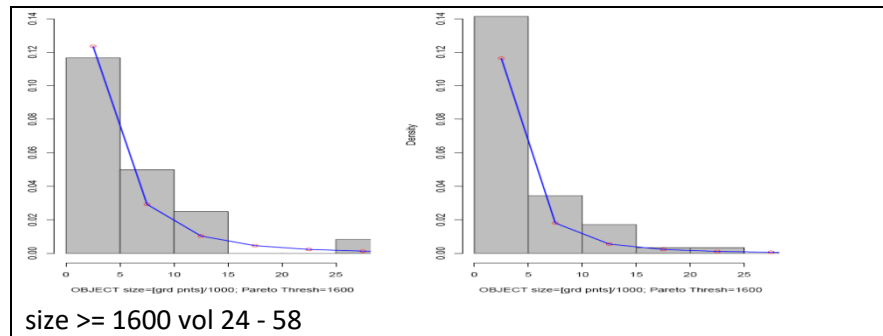


Figure 8. Cold period. Histograms and Pareto distribution approximation of object sizes in precipitation fields in Smolensk locator observations (RUDL, left column) and in forecasts (STEPS-60, right column) for 60 min. The Pareto threshold is (from top to bottom) 625, 900, 1225 and 1600 field points. Dimensions are given in size/1000 scale. **Blue lines!!!**

The full set of situations that generated the peaks above the thresholds is presented in Table 10 for a threshold of 625 points; as the threshold increases, the sample is made up of the sizes of the same objects. In the presented set, there are two situations, in which the maximum object turned out to be the only one: at numbers 32 and 33, while at number 33 the situation had a duration of 60 minutes, but the generated object had a size of 13044 points, which is equivalent to a square with a side of about 230 km (almost a quarter from the field of view of the radar). However, the largest object appeared in situation 18, which lasted 1280 min (almost a day), while in 128 fields of radar observations there were a total of 125 objects with a size of at least 632 points. It is quite possible that in a real synoptic situation, the total number of objects constructed by spatial averaging at 9 grid nodes and an isohyet of 1 mm constituted one system of frontal-type precipitation.

Table 10. Spatial, statistical and temporal characteristics of *situations* with objects with a size of at least 625 points in a two-kilometer resolution grid observed in the Smolensk radar zone during the cold period November 2017 - March 2018 The objects are marked with a 1 mm/h isoline after spatial averaging with a radius of 9 points.

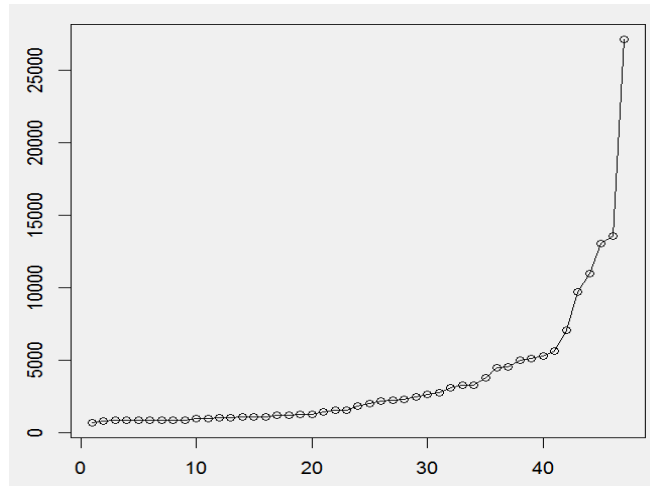


Figure 9 . Graph of the variational series, made up of a set of forty-seven extrema in Table 10 (column max).

3.3.3. Integral Quality Estimates Using Distribution Parameters

Let's compare the parameters of the Pareto distribution for maximum objects in situations of observations and forecasts based on the data of the Kursk (RAKU) radar (Table 11). Consider, for example, the result of a forecast for 60 minutes. For observations in the warm period, the results are as follows: 87 situations were identified; scale parameter estimate and standard error are 1956 and 472.9, respectively; shape parameter estimate and standard error are 0.428 and 0.212. Recall that the boundaries of the confidence interval are determined in a standard way (estimate $\pm 1.96 \cdot$ error). Similar numbers for the forecast for 60 minutes are as follows: 119 situations are identified; scale parameter estimate and standard error are 1979 and 391.1, respectively; shape parameter estimate and standard error are 0.413 and 0.180. Let's write lower and upper limits of the confidence intervals like (L_1, U_1) and (L_2, U_2) , respectively. The intersection ratio (IR), visually obvious, is determined as follows:

$$IR = (\min(U_1, U_2) - \max(L_1, L_2)) / (\max(U_1, U_2) - \min(L_1, L_2)),$$

when $DP < 0$, $DP = 0$ is assigned. For the scale parameter in the selected case, $IR = 0.83$, for the shape parameter, $IR = 0.85$. In the cold period of 2017-2018, the following results were obtained: for the scale parameter $IR = 0.50$, and for the shape parameter $IR = 0.68$.

Table 11. Number of situations (# situations), estimates of scale parameters (scale), shape (shape) of the Pareto distribution of object sizes with thresholds of 625, 1225, 900 and 1600 grid points, the intersection ratio (*intersect*) of confidence intervals of parameter estimates. Objects are extracted from situations in series of fields of observations (RAKU) and forecasts (STEPS) with lead times of 30, 60, 90 and 120 min. Observation and forecast data refer to two periods: warm - May-September 2017 (yellow) and cold - November 2017 - March 2018 (green).

Lead time 30 min							
Threshold		WARM PERIOD			COLD PERIOD		
		scale	shape	#situations	scale	shape	#situations
625	RAKU	1956	0.428	87	2617	0.294	59
	STEPS	1859	0.420	117	2105	0.410	87
	<i>intersect</i>	0.80	0.84		0.75	0.78	
900	RAKU	2497	0.362	69	3415	0.000	51
	STEPS	2504	0.321	94	2833	0.324	70
	<i>intersect</i>	0.74	0.74		0.68	0.47	
Lead time 60 min							
625	STEPS	1979	0.413	119	1875	0.457	106
	<i>intersect</i>	0.83	0.85		0.50	0.68	
900	STEPS	2621	0.320	97	2891	0.320	81
	<i>intersect</i>	0.77	0.76		0.63	0.44	
Lead time 90 min							
625	STEPS	1958	0.409	121	1342	0.558	129
	<i>intersect</i>	0.83	0.83		0.23	0.54	
900	STEPS	2268	0.368	103	2248	0.395	96
	<i>intersect</i>	0.73	0.76		0.38	0.33	
Lead time 120 min							
625	STEPS	2062	0.382	123	1338	0.548	142
	<i>intersect</i>	0.79	0.80		0.21	0.54	
900	STEPS	2250	0.364	107	1837	0.459	112
	<i>intersect</i>	0.68	0.72		0.20	0.23	

The question arises about the statistical significance of the obtained numbers. The IR distribution is rather complicated (an attempt to derive a distribution formula for the Gaussian case is given in [Ants 1990]), although mathematically the problem is reduced to estimating the probability of a joint event. However, according to the percentage of intersection, it can be argued that the quality of forecasting objects of a significant size for a period of 60 minutes in the cold period (2017-2018) according to the Kursk radar exceeds the quality of forecasting in the warm period. We note that only the data of this radar differ in this property (see Table 12 for comparison).

Let's continue generalizing the verification results for all radars used in the 2017-2018 tests. We choose Pareto thresholds of 625 and 900 points as the most informative, which provide a reliable basis for data analysis. Note that for a more stable estimate of the Pareto parameters for the selected situations, only the samples with at least 20 precipitation situations were allowed. For both periods, the Vnukovo (RAVN , in winter) and Valdai (RUWJ) radars did not satisfy this condition to a large extent - these data were completely excluded. In the set of cold period fields for the Kostroma (RUDK) and Nizhny Novgorod (RUDN) locators , the condition turned out to be critical only at a threshold of 1600 points.

Let's recall the interpretation of the parameters of the Pareto distribution of maximum values over the selected threshold. The scale determines the value of the GPD at the zero point ($GPD(0)=1/\text{scale}$). The larger the scale, the smaller the GPD value at zero and the lower the probability of

having an object with a size close to the Pareto threshold. Thus, we can say that the scale characterizes more clearly the probabilistic features of objects with sizes closer to the Pareto threshold. The shape parameter is more important in GPD analysis, since it determines the characteristic of the "tail of the distribution", that is, the probability of the largest areas of precipitation. The convenience of the shape parameter lies in the meaning of the parameter sign: the negative sign, the zero form, and the positive sign indicate, respectively, the beta distribution, the exponential distribution, and the actual Pareto distribution. The value of the positive shape parameter is also significant: it ensures the existence of moments of different orders for a distribution of this type.

Let us consider the intersection ratio (IR) of confidence intervals for estimates of the scale and shape parameters (Table 12). Let's choose a level of "failure", e.g., intersect <50%, and mark it in red. There are various features in the behavior of the IR depending on the lead time and the Pareto threshold, some systematic, some random, but here we will indicate only the most noticeable of them, based on the values highlighted in red. Prediction of precipitation areas is better in the cold period according to the RATL and RAVO radars, and in the warm period, according to the RAKU , RUDB and RUDL radars .

Table 12. Intersection ratios of confidence intervals (intersect) of the estimate of the scale parameter and the shape parameter for Pareto thresholds of 625, 900, 1225, and 1600 points, for lead times of 30, 60, 90, and 120 min, for the warm and cold periods of 2017–2018 Values less than 50 are marked in red. Cases of insufficient number of situations in observations are marked with an asterisk.

RADAR		IR (%) SCALE			
		warm period		cold period	
	threshold\ lead time	625	900	625	900
RAKU	30	80	74	75	68
	60	83	77	50	63
	90	83	73	23	38
	120	79	68	21	20
RATL	30	39	27	62	78
	60	48	23	52	55
	90	35	17	47	53
	120	34	19	50	62
RAVO	30	54	38	70	76
	60	58	37	64	74
	90	56	46	61	63
	120	52	35	50	64
RUDB	30	92	88	75	78
	60	87	90	48	67
	90	81	94	46	56

IR (%) SHAPE			
warm period		cold period	
625	900	625	900
84 (+ +)	74 (+ +)	78 (+ +)	47 (0 +)
85 (+ +)	76 (+ +)	68 (+ +)	44 (0 +)
83 (+ +)	76 (+ +)	54 (+ +)	33 (0 +)
80 (+ +)	72 (+ +)	54 (+ +)	23 (0 +)
74 (+ +)	41 (0 +)	78 (+ +)	87 (+ +)
72 (+ +)	36 (0 +)	69 (+ +)	71 (+ +)
66 (+ +)	32 (0 +)	67 (+ +)	69 (+ +)
65 (+ +)	34 (0 +)	68 (+ +)	73 (+ +)
78 (+ +)	40 (0 +)	76 (+ +)	71 (+ +)
76 (+ +)	36 (0 +)	72 (+ +)	67 (+ +)
77 (+ +)	38 (0 +)	66 (+ +)	63 (+ +)
69 (+ +)	34 (0 +)	64 (+ +)	60 (+ +)
93 (+ +)	91 (+ +)	72 (+ +)	72 (+ +)
91 (+ +)	92 (+ +)	60 (+ +)	64 (+ +)
92 (+ +)	94 (+ +)	56 (+ +)	57 (+ +)

	120	88	92	44	56	89 (+ +)	92 (+ +)	55 (+ +)	56 (+ +)
RUDK	30	78	85	69	73	80 (+ +)	79 (+ +)	75 (+ +)	70 (+ +)
	60	76	80	54	68	75 (+ +)	74 (+ +)	64 (+ +)	64 (+ +)
	90	80	82	50	60	76 (+ +)	74 (+ +)	60 (+ +)	57 (+ +)
	120	72	80	52	50	74 (+ +)	73 (+ +)	56 (+ +)	53 (+ +)
RUDL	30	75	63	47	52	80 (+ +)	76 (+ +)	72 (+ +)	71 (+ +)
	60	73	64	32	47	76 (+ +)	73 (+ +)	64 (+ +)	63 (+ +)
	90	68	57	40	45	73 (+ +)	69 (+ +)	62 (+ +)	61 (+ +)
	120	71	66	41	42	74 (+ +)	70 (+ +)	59 (+ +)	57 (+ +)
RUDN	30	78	85	87	70	89 (0 0)	80 (0 0)	87 (+ +)	89 (+ +)
	60	52	86	57	71	38 (0 +)	78 (0 0)	69 (+ +)	71 (+ +)
	90	38	83	54	58	30 (0 +)	79 (0 0)	65 (+ +)	63 (+ +)
	120	31	84	57	54	27 (0 +)	80 (0 0)	64 (+ +)	59 (+ +)

Let's analyze the confidence intervals of the shape parameter estimate (shape, Table 12). Let's add the sign (+/-) depending on the sign of the shape for observations/predictions, with the sign 0 - for the shape interval $[-0.1, 0.1]$. It is desirable that the forecasting system preserves the sign of the shape parameter. According to the table, there is a general dependence of the sign of the shape on the Pareto threshold. So, in the warm period, the parameter generally decreases when moving from a threshold of 625 to a threshold of 900. We add that when passing to the rejected thresholds 1225 and 1600, the shape parameter estimate goes to zero and even to negative. This boundary between the thresholds provides further evidence that areas larger than 625 points (approximately 50×50 km) are most suitable for Pareto analysis in both observations and forecasts. Indeed, for all radars, except for RUDN (Nizhny Novgorod), the sign (++) for all lead times and for both periods indicates that both observations and forecasts fit the Pareto distribution. At the same time, the value of the intersection ratio (at the threshold of 625) is noticeably higher in the warm period for the RAKU (Kursk), RAVO (Voeykovo), RUDB (Bryansk), and RUDL (Smolensk) radars. The picture is different for the RATL (Tula) and RUDN (Nizhny Novgorod) radars, which are located to the east of the above, and where the quality of reproduction of vast contiguous precipitation areas is higher.

Conclusions

The study considers the problems of modeling of extreme values on the example of contiguous precipitation areas observed and predicted by the precipitation nowcasting system in the coverage areas of DMRL-S radars deployed in the Central Federal District of Russia. *Precipitation areas* were converted into *objects* using spatial averaging and the isoline of 1 mm/h. The sets of such *sizes* (or areas) of objects exceeding certain threshold values were formed so that they at least partially satisfy the conditions of physical (and at the same time statistical) independence for applying the extreme value theory (EVT). The model of "peaks above the threshold" described by the generalized Pareto distribution is chosen as the basic model of extreme values.

All the main computational procedures were performed using the tools and graphical representation available in the R language. Objects were selected using the mathematical module `FeatureFinder()` of the `SpatialVx` library. To estimate the distribution parameters, we selected objects with sizes of at least 25×25 , 30×30 , 35×35 , and 40×40 points in a two-kilometer grid. The generalized Pareto distribution was used with fixed location thresholds (Pareto thresholds) equal to the selected object sizes (625, 900, 1225 and 1600 points). The parameters were estimated using 1) maximum likelihood methods, 2) maximum likelihood, 3) L-moments, and 3) Bayes with stochastic Markov chain modeling.

The conclusions are based on the generalized maximum likelihood method and two Pareto thresholds, 625 and 900 points. The output standard errors of the estimates are used to construct 95% confidence intervals (CI) and to subsequently compare estimates of the scale and shape parameters based on the intersection ratio (IR). Particular attention is paid to the shape parameter, the positivity of which ("Pareto" distribution of extrema) indicates the presence of a heavy tail in the distribution: the larger the value of the shape parameter, the heavier the tail and the more problematic the existence of distribution moments.

It is shown that with increasing threshold (from 625 to 1600) the shape parameter tends to change sign from positive to zero and, in rare cases, to negative. The zero sign in observations and forecasts at a threshold of 625 points was observed for only one radar (RUDN) during the warm period. Negative estimates of the shape parameter are even rarer; at the threshold of 625 points, such cases are completely absent.

Assuming the IR of 50% or more as an acceptable error, two conclusions can be drawn.

First, the precipitation nowcasting system better predicts objects of extreme size in the cold season. The number of pairs (++) in the warm period according to the table 12 is about half of the cases, and in the cold - about 75%.

Secondly, the precipitation nowcasting system most accurately reproduces the Pareto distribution of precipitation areas in the warm period - in the coverage areas of the RAKU (Kursk), RAVO (Voeykovo), RUDB (Bryansk), RUDL (Smolensk) radars, and in the cold period - in the coverage areas to the east of the RATL (Tula) and RUDN (Nizhny Novgorod) radars.

The last conclusion from the work done can be attributed to the methodology: the extreme value theory is quite applicable to such objects of analysis and short-term forecasting as significant contiguous precipitation areas only with a clear understanding of the theoretical prerequisites and using suitable statistical methods and reliable data processing tools. Otherwise, the results obtained may be useless, accidental, or even harmful.

References

1. *Abdullaev S.M.* Life cycle of mesoscale convective systems: concept, climatology and forecast: Diss. d.g.s. M, 2010, 408 p. [In Russian]
2. *Alibegova Zh.D.* Spatial structure of the field of total precipitation averaged over 15-minute intervals // Proceedings of the GGO. 1972. issue. 280. p. 195-208. [In Russian]
3. *Alibegova Zh.D.* Spatio-temporal structure of liquid precipitation fields L.: Gidrometizdat, 1985. 224 p. [In Russian]
4. *Belyaev Yu.K.* Limit theorems for thinning flows, *Teor. Veroyatnost. and its applications*, 1963, volume 8, issue 2, 175–184. [In Russian]
5. *Bolshev L.N., Smirnov N.V.* Tables of mathematical statistics. M.: Nauka, 1983. 416 p. [In Russian]
6. *Bundel A.Yu., Muravyov A.V., Olkhovaya E.D.* Review of Spatial Verification Methods and Their Application for Ensemble Forecasts // *Hydrometeorological Research and Forecasting*. 2021. No. 4 (382). pp. 30–49. [In Russian]
7. *WMO-168.* Guide to hydrological practice. Volume I I . Water resources management and practice of application of hydrological methods. Ed.6, 2012 (2009). Chapter 5. Extreme Value Analysis. II -5.1 - II - 5.58.
8. *Veltishchev N.F., Stepanenko V.M.* Mesometeorological processes. M.: Faculty of Geography of Moscow State University, 2006, 100 p. [In Russian]
9. *WMO-168.* Guide to hydrological practice. Volume II . Ch. 5. Analysis of extreme values. 2009. II.5-1- II.5-63.
10. *Temporary guidelines* on the use of information from the Doppler meteorological radar DMRL-S in synoptic practice. M.: Roshydromet, Second edition, 2017, 121 p. [In Russian]
11. *Galambosh J.* Asymptotic theory of extremal order statistics. Moscow : Nauka , 1984, 304 p. [In Russian]
12. *Galambosh Ya.* On the development of the mathematical theory of extrema over the past half century // *Theory of Probability and Its Applications*, 1994. V. 39. No. 2. P. 272–293. [In Russian]
13. *Gnedenko B.V.* Limit theorems for the maximum term of a variational series. *RAN USSR*, 32, No. 1, 7. 1941. [In Russian]
14. *Gnedenko B.V.* Probability course. Ed. 6th. M.: Nauka, 1988, 448 p. [In Russian]
15. *Gnedenko B.V., Fryer B.* Some remarks on one work of N.N. Kovalenko. *Lithuanian mathematical collection*. 1969, IX , 3, 463-470. [In Russian]
16. *Goldaeva A.A.* Extremal indices and clusters in linear recurrent stochastic sequences // *Theory of Probability and its Applications*. 2013. V. 58. No. 4. S. 795–804. [In Russian]
17. *Gorshenin A.K.* On some mathematical and software methods for constructing structural models of information flows. *Informat. i ee prilozheniya*, 2017, 11, issue 1, 58-68. [In Russian]
18. *Gorshenin A.K.* Semiparametric methods for the analysis of heterogeneous data and their application in problems of mathematical modeling. Diss. Doctor of Physics and Mathematics, 2021. [In Russian]
19. , *Korolev V.Yu.* Determination of the extremeness of precipitation volumes based on the modified method of exceeding the threshold value .*Informat. and its applications*, 2018, 12, issue 4, 16-24. [In Russian]
20. *Gumbel E.* Statistics of extreme values. M.: Mir, 1965, 453 p. [In Russian]
21. *Gumbel E.* Statistical theory of extreme values (main results) / Introduction to the theory of order statistics. M.: Fazis, 1970. S. 61–93. [In Russian]
22. *De Groot M.* Optimal statistical solutions. M., Mir, 1974, 492 p.
23. *Zolotarev V.M.* One-dimensional stable distributions. Moscow: Nauka, 1983. [In Russian]
24. *Isaev A.A.* Precipitation. Part II . Mesostructure of liquid precipitation fields. M.: Faculty of Geography of Moscow State University, 2001. 100 p. [In Russian]
25. *Kiktev D.B., Muravyov A.V., Smirnov A.V.* Nowcasting of meteorological parameters and hazardous phenomena: implementation experience and development prospects. *Hydrometeorological research and forecasts*. 2019, 4(374). Anniversary collection "90 years of the Hydrometeorological Center of Russia", pp. 92-111. [In Russian]
26. *Kiktev D.B., Muravyov A.V., Smirnov A.V.* Prediction of Precipitation with a Multiplicative Cascade Model: A Radar Nowcasting Experience. *Turbulence, atmospheric and climate dynamics*. International

conference dedicated to the 100th anniversary of Academician A.M. Obukhova, Moscow, May 16-18, 2018. Abstracts, p. 27. [In Russian]

27. Kolmogorov A.N. Basic concepts of probability theory. M.: Nauka, 1974, 119 p. [In Russian]

28. Kolmogorov A.N., Fomin S.V. Elements of the theory of functions and functional analysis. 4th ed. M.: Nauka, 1976, 544 p. [In Russian]

29. Kramer G. Mathematical methods of statistics. M.: Mir, 1975, 648 p. [In Russian]

30. Lebedev A.V. Fundamentals of the stochastic theory of extrema. M.: LENAND, 2018. 104 p. [In Russian]

31. Lenskaya O.Yu. Mesoscale organization and evolution of sedimentary systems in southern Brazil. Diss. Ph.D. M. 2006, 220 p. [In Russian]

32. Leadbetter M., X. Rotsen, G. Lidgren. Extrema of random sequences and processes. M.: "Mir", 1989. [In Russian]

33. Markovich N.M. Statistical analysis of clusters of extreme values and its application // Proceedings of the 9th International Conference "Identification of Systems and Control Problems" (SICPRO'2012, Moscow). M.: IPU RAN, 2012. T 1. S. 1024-1034. [In Russian]

34. Muravyov A.V., Kiktev D.B., Smirnov A.V. Operational technology of rainfall nowcasting based on radar data and verification results for the warm period of the year (May-September 2017) // Hydrometeorological Research and Forecasting. 2018. No. 1 (367). C. 6-38. [In Russian]

35. Muravyov A.V., Kiktev D.B., Smirnov A.V., Zaichenko M.Yu. Operational technology of rainfall nowcasting based on radar data and comparative results of point verification for warm and cold periods of the year // Hydrometeorological Research and Forecasting. 2019. No. 2 (372). C. 12-40. [In Russian]

36. Smirnov A.V. Improved technology of precipitation radar nowcasting and verification results in the warm season (May - September 2020) // Results of testing new and improved technologies, models and methods of hydrometeorological forecasts. 2021. Information collection No. 48. S. 13-41. [In Russian]

37. Muraviev A.V., Kiktev D.B., Smirnov A.V. Comparative verification of an improved system of precipitation radar nowcasting taking into account gaps and with various sampling methods (according to the results of tests in the warm period of the year May-September 2017 and 2020) // Results of testing new and improved technologies, models and methods of hydrometeorological forecasts. 2022. Information Collection No. 49. 3-56. [In Russian]

38. Muraviev A.V., Bundel A.Yu., Kiktev D.B., Smirnov A.V. Experience in spatial verification of precipitation radar nowcasting: identification and statistics of objects, situations and conditional samples. Hydrometeorological Research and Forecasting, 2 (384), 2022, p. 6-52. [In Russian]

39. Muraviev A.V., Bundel A.Yu., Kiktev D.B., Smirnov A.V., Verification of radar nowcasting of vast precipitation areas using generalized Pareto distribution, Part I: Theory elements and methods of estimation of parameters. Hydrometeorological Research and Forecasting. In press. [In Russian]

40. Nazarenko K.M. On a new method for modeling multidimensional extreme values based on the threshold approach. Bulletin of RUDN University, Series Mathematics. Informatics. Physics. No. 2. 2008. p. 30-38. [In Russian]

41. Rodionov I.V. Probabilistic and statistical analysis of extrema of discrete stochastic systems. Diss. Doctor of Physics and Mathematics, 2021.

42. Seneta E. Correctly changing functions. M.: Nauka, 1985, 144 p. [In Russian]

43. Smirnov N.V. Limit distribution laws for members of a variational series, Tr. MIAN USSR, 1949, volume 25, 3-60 [In Russian]

44. Feller W. Introduction to probability theory and its applications. T.1., T.2. M.: Mir, 1984. [In Russian]

45. Shetin E.Yu. Mathematical theory of extreme values in modeling and evaluating financial risks. Financial Informatics, 2008 1(1), 63-70 [In Russian]

46. Akaike, H. (1974). A new look at the statistical model identification. IEEE Transactions on Automatic Control, 19, 716-723.

47. Balkema A., De Haan L. Residual life time at great age, Ann. Probab. 2(5): 792-804 (October, 1974).

48. Beirlant J., Goegebeur Y., Teugels J., Segers J., De Waal D., Ferro C. Statistics of Extremes. Theory and Applications. John Wiley & Sons Ltd, England. 2004, 504 p.

49. Bowler N., Pierce C., Seed A. STEPS: A probabilistic forecast precipitation scheme which merges an extrapolation nowcast with downscaled NWP // QJR Meteorol. soc. 2006 Vol. 132. P. 2127-2155.

50. Coles S. G. (2001). An Introduction to Statistical Modeling of Extreme Values. Springer Series in Statistics. Springer-Verlag London Ltd., London.
51. Davis CA, Brown BG, Bullock RG 2006: Object-based verification of precipitation forecasts, Part I: Methodology and application to mesoscale rain areas. *Monthly Weather Review*, 134, 1772-1784.
52. Davis CA, Brown BG, Bullock RG 2006: Object-based verification of precipitation forecasts, Part II: Application to convective rain systems. *Monthly Weather Review*, 134, 1785-1795.
53. Davison AC, Smith RL Models for exceedances over high thresholds. *JR Statist. Soc. B*(1990), 52, no. 3, 393-442.
54. De Haan L, Resnik S . On asymptotic normality of the Hill estimator, *Stochast. Mod.*, 12. R. 699 - 724, 1996.
55. De Haan L. On regular variation and its application to weak convergence of sample extremes. *Math. Center Tracts*, 1970, Vol. 32, Amsterdam.
56. De Haan L., Ferreira A. *Extreme value theory. Introduction*. Springer, 2006.
57. Degen M., Embrechts P., Lammbrigger DD . (2007): The quantitative modeling of operational risk: between g-and-h and EVT. *Astin Bulletin* 37(2), 265-291
58. Embrechts P., Klueppelberg C., Mikosch T . *Modeling extreme events for insurance and finance*. Springer, 2003.
59. *Extremes and Integrated Risk Management* . Ed. Embrechts, P. (2000). Risk Books, London. 397 r .
60. *Extreme Events in Finance - A Handbook of Extreme Value Theory and its Applications*. Ed . Longin F. _ Wiley , 2017, 640p .
61. Ferreira A., De Haan L. *Extreme value theory. Introduction*. Springer Series in Operations Research and Financial Engineering. NY: Springer, (2006).
62. Ferro CAT Statistical Methods for Clusters of Extreme Values. Submitted for the degree of Doctor of Philosophy at Lancaster University, September 2003 (Thesis). <http://empslocal.ex.ac.uk/people/staff/ferro/Publications/Thesis/thesis.pdf>
63. Ferro, CAT and Segers, J . Inference for clusters of extreme values. *Journal of the Royal Statistical Society* . 2003. B, 65, 545–556.
64. Fisher RA, Tippett LHC Limiting forms of the frequency distribution of the largest or smallest member of a sample // *Proc. Camb. Phil. soc.* 1928. V. 24. P. 180–190.
65. Galambos J., Seneta E. (1973), Regular varying sequences, *Proc. amer. Mat. Soc.*, 41(1): 110–116
66. Geyer CJ Introduction to Markov Chain Monte Carlo. (In: *Handbook. Markov Chain Monte Carlo*, Eds Brooks S., Gelman A., Jones G., Meng X.-L. 2011, Chapman and Hall/CRC.) p. 1-77.
67. Gilleland E . 2020, <https://cran.r-project.org/web/packages/SpatialVx/SpatialVx.pdf>
68. Gilleland E . Bootstrap Methods for Statistical Inference. Part I: Comparative Forecast Verification for Continuous Variables. *J. Atm. ocean. Technol.*, 2020, vol.37, 2117-2134. doi: 10.1175 /JTECH-D-20-0070.1
69. Gilleland E . Bootstrap Methods for Statistical Inference. Part II: Extreme-Value Analysis. *Jour. Atm. ocean. Technol.*, 2020, vol.37, 2135-2144. doi: 10.1175/JTECH-D-20-0070.1
70. Gilleland E., Katz RW . extRemes 2.0: An Extreme Value Analysis Package in R. *Journal of Statistical Software*. August 2016, Volume 72, Issue 8. doi:10.18637/jss.v072.i08 .
71. Gnedenko BV Sur la distribution limite du terme maximum d' une serie aleatoire // *Ann . Math.* 1943. V. 44. No. 3. P. 423–453.
72. Hill BM (1975) A simple general approach to inference about the tail of a distribution. *Annals of Statistics* 3 , 1163–1174.
73. Jenkinson AF The frequency distribution o the annual maximum (or minimum) values of meteorological elements. *Quart.J. Roy. meteor. Soc.*1955, Vol. 81, 348, 158-171.
74. Karamata J. Neuer Beweis und Verallgemeinerung der Tauberschen Sätze, welche die Laplace'sche und Stieltjes'sche Transformationen betreffen, *J. Reine Angew. Math.* 164 (1931) 27–39.
75. Karamata J. Sur un mode de croissance régulière des fonctions, *Mathematica (Cluj)* 4 (1930) 38–53 .
76. Leadbetter MR On extreme values in stationary sequences. — *Z. Wahrsch. verw. Geb.*, 1974, v. 28, p. 289-303
77. Leadbetter MR, Lindren G., Rootzen H . *Extremes and related properties of random sequences and processes*. New York: Springer; 1983.
78. Loynes RM Extreme values in uniformly stationary mixing stochastic processes. — *Ann. Math. Statist.*, 1965, v. 36, p. 993-999.

79. Markovich NM . Clusters of extremes: modeling and examples // Extremes. 2017.V20. No. 3. P. 519–538.
80. Martins, E. S. and Stedinger, J. R. (2000). Generalized maximum likelihood extreme value quantile estimators for hydrological data. *Water Resources Research*, 36(3), 737–744.
81. Martins, E. S. and Stedinger, J. R. (2001). Generalized maximum likelihood Pareto-Poisson estimators for partial duration series. *Water Resources Research*, 37(10), 2551–2557.
82. McNeil, A., Frey, R., and Embrechts, P. (2005). *Quantitative Risk Management: Concepts, Techniques, and Tools*. Princeton University Press, 720 p.
83. Mendez FJ, Menendez M., Luceno A., Losada IJ . Estimation of the long-term variability of extreme significant wave height using a time-dependent Peak Over Threshold (POT) model. *J. _ Geophys . Research* , Vol . 111, C 07024, 2006, 1 - 13.
84. Mises von, R. (1936). La distribution de la plus grande de n valeurs. *Rev. Math. Union Interbalcanique* 1, 141–160. Reproduced in *Selected Papers of Richard von Mises*, Amer. Math. soc. 2 (1964), 271–294.
85. Nešlehová J., Embrechts P., Chavez-Demoulin V. (2006). Infinite mean models and LDA for operational risk. *Journal of Operational Risk* 1 (1), 3–25.
86. Novak S.Y. (2011) *Extreme value methods with applications to finance*. – London: Chapman & Hall/CRC Press.
87. Pancheva E . Limit theorems for extreme order statistics under nonlinear normalization. *Lect. Notes Math.*, 1984, v. 1155, p. 284-309.
88. Pancheva E . Convergence to type theorems with monotone mappings. *Theory probable . and her appl .*, 1993, vol . XXXVIII, c . 3, p . 679-684.
89. Papalexiou SM Koutsoyiannis D. 2013. Battle of extreme value distributions: a global survey an extreme daily rainfall. *water resource. Res.* 49(1). 187-201.
90. Pickands J. Statistical inference using extreme order statistics, *Annals of Statistics*, 3, 119–131, 1975.
91. Raoult J.-P., Worms R . (2003) Rate of convergence for the generalized Pareto approximation of the excesses. *Advances of Applied Probability* 35, 1007–1027.
92. Reiss R.-D., Thomas M. *Statistical Analysis of Extreme Values with Applications to Insurance, Finance, Hydrology and Other Fields*. 3d Ed. Berlin, Birkhauser Verlag, 2007, 516 p.
93. Resnick SI *Extreme values, regular variation and point processes*. Springer, 1987.
94. Schwarz, G.E. (1978). Estimating the dimension of a model. *Annals of Statistics* , 6 , 461–464.
95. Seneta E. *Regularly Varying Functions*. *Lecture Notes in Mathematics*, Vol. 508, Springer-Verlag, Berlin-New York, 1976.
96. Tawn J. A. (1988). An extreme value theory model for dependent observations. *J. Hydrol.*, 101, 227-250.
97. Tawn JA, Shooter R., Towe R., Lamb R. (2018). Modeling spatial extreme events with environmental applications. *Spatial Statistics*, 28:39-58.
98. Wilks D.S. *Statistical methods in the atmospheric sciences*. 4th Ed. 2019 Elsevier. 816p.
99. WMO -No.168 . *Guide to Hydrological Practices, Volume II: Management of Water Resources and Applications of Hydrological Practices*. From h.5. Extreme value analysis . 2009 , II.5-1 - II.5-59 . _