

A. B. Ivanovsky
RosNeft – Inform Ltd. Krasnoyarsk Subsidiary, Russia, Krasnoyarsk

V. V. Shishov
Krasnoyarsk State Institute of Trade and Economy, Russia, Krasnoyarsk

A PARAMETERIZATION ALGORITHM FOR THE VAGANOV–SHASHKIN MODEL OF SEASONAL GROWTH AND TREE-RING FORMATION*

In order to simulate the Vaganov–Shashkin model for seasonal growth and tree-ring formation, a solution algorithm for the parameterization problem of the model is being proposed in cases, when a modulation is possible. The algorithm is realized as dll-library (or as a text file), tested on extensive data. A concept of difference in criterion between the actual tree-ring chronology and its model is introduced. Two new difference criteria are developed.

Keywords: the Vaganov–Shashkin model, tree-ring chronology, parameterization algorithm, difference criterion, optimal model parameters.

The Vaganov- Shashkin model of seasonal growth and tree-ring formation [1–3] (hereinafter referred to as VS-model) describes the influence of climatic conditions on the cellular structure of annual rings. The main destination of the model consists in using it as a tool permitting any individual or generalized tree-ring chronology (TRC) during a set of years with available meteorological station data (see below description of the VS-model input): 1) extracting the climate-driven component of the TRC under consideration; 2) pointing, if modeling quality of the involved TRC by the VS-model is satisfactory, for arbitrary day of the given years set, what of two factors – air temperature or precipitation – had limited the growth of woody plant which corresponds to the TRC. TRC is a time series of values of some numerical characteristic of tree/trees of this TRC [4].

The main destination of the VS-model, in particular, is that it makes possible to use it as a quality tester for individual and generalized TRCs which will be used as mediators, carrying information about climatic data [2]. A TRC is considered to be suitable for using for the purpose of a mediator for carrying climatic information if and only if its climate-driven component had been extracted by the VS-model conforms with the TRC satisfactorily enough.

The VS-model is a deterministic dynamic simulation model. Its input data consist of two blocks. The first block is the climatic data of daily resolution at air temperature, precipitation, and doses of solar radiation coming down to the Earth's surface. The data of this block can refer to an arbitrary set of years, which does not have to be continuous. However, climatic data per year must not contain missing values.

The second block is a set of the VS-model's parameter values. The model has 42 parameters. All parameters, except two, are real single-valued variables and three parameters are integer single-valued variables. Two parameters are vectors of equal dimension. One of the vectors contains real single-valued variables, the second – boolean single-valued variables. The dimension of the vectors is a parameter of the VS-model.

The value of each parameter provides the VS-model with information either on the actual tree or on its site. The VS-model has 10 options. A set of option values defines the VS-model variant which will be used for modeling. The options of the VS-model define a collection of its modules being run under modeling, calculating accuracy under modeling, etc.

The output of the VS-model consists of two blocks. The data of the first block consists of numerical characteristics reflecting the dynamics in time, the time step of whose tracking does not exceed one day – for some processes taking place in the modeled tree and its modeled site (soil moisture, value of tree transpiration, number of cambial cells, etc). The contents of this block depend on the used set of the VS-model option values.

The second output block of the VS-model is a model of TRC. The VS-model supposes the actual individual TRC for being modeled to be a time series of values either of width of a formed annual ring or of number of wood cells in such a ring. The VS-model supposes all individual TRC are used for the construction of an actual generalized TRC to be such a one described above. For each year of the VS-model, the climatic input of the second block contains the value of modeled TRC.

This value does not have any units if a generalized TRC has been modeled. If an individual TRC has been modeled, this value either does not have any units or is number of wood cells in an annual ring (variants, taking place are determined by the used set of the VS-model options values).

A parameterization problem and approach to its solution for the VS-model. In the field of mathematical modeling the term “parameterization” is usually understood as either an activity on describing some process/phenomenon by means of a finite number of parameters, i. e. the creation of a parametric mathematical model of this process/phenomenon, or choosing certain values of parameters of a parametric mathematical model that were already created. In the article the term is being understood according to the second interpretation.

*This work has been supported by Russian Foundation of Basic Research (Project № 09-05-00900-a).

To process the VS-model it is necessary to have certain numerical values of its parameters. Sets of these values must have the necessary internal structures that are specified by interrelations between parameters of the VS-model and by ranges of parameter values. Meaning, many parameters, the VS-model structure, semantics of its output components, and specific context in which modeling is performed determine the mentioned interrelations, and value ranges. As a rule, the checking a set of values of the VS-model parameters for the presence of the necessary internal structure in it requires an analysis of the VS-model output that corresponds to this set.

In an attempt to set certain values to the model's parameters to realize the modeling, it often happens that the available information is not enough for the identification of exact values for some parameters. For such parameters it is possible to provide only a qualitative assessment of their values, for example, by means of determining the boundaries in which the values of these parameters lie. In addition to a number of properties of some the VS-model's parameters, it complicates the empirical measurement of the values and makes the exact value conception for these parameters meaningless. It has as a consequence in a situation when for each parameter of the VS-model, the range of its values was known, and for some parameters their specific numerical values are known.

The parameterization problem for the VS-model is a problem of choosing values of those parameters, for which only the ranges of their values are known. The choice has to be realized within these ranges in such a way that the derived set of values of all the VS-model parameters has the necessary internal structure. How do we realize this choice? What should one be guided by to realize the choice? These two questions are the essence of the parameterization problem. The parameterization problem of an existing mathematical model is a problem of choosing the model input data in conditions of insufficient information.

There are different approaches used in the practice of solving the parameterization problem of a parametric mathematical model [5]. A solution to the parameterization problem for the VS-model, presented below, uses the existing arbitrariness in choosing values of parameters, for which information is necessary for determining their exact values is absent, purposively. Values of the parameters are selected, which conform to the researcher's objective formalized in a form of reaching the optimum of an objective function in the optimal, over the domain of definition of the objective function, way (conform in the optimal way). The proposed solution to the parameterization problem for the VS-model can be applied only in cases when entity for being modeled is available, in particular, under using the VS-model according to its main destination.

A concept of difference criterion between actual TRC and its model. Introducing the concept "difference criterion between actual TRC and its model", being significant out of the context of the VS-model

parameterization problem, to state the solution being proposed in this article to the latter problem. The actual TRC is a TRC for being modeled. The difference criterion is a real single-valued nonnegative function of two arguments, denote it by $DC(trCr, mCr)$, that is defined over a set of pairs of real vectors of equal dimension with nonnegative components. The dimension of the vectors is the power of set Yrs . The set Yrs is the intersection of two sets – set of years, which the modeling has been performed for, and set of those years, which values of the actual TRC are available for.

The first argument is that $trCr$ must be a vector of values, relating to the years Yrs , of the actual TRC, which are situated in the vector in increasing order of the years. The second argument must be mCr a vector of respective values of the TRC modeled by the VS-model.

A value of the difference criterion characterizes the remoteness of the TRC model, obtained by means of the VS-model, from this actual TRC. Given a fixed actual TRC, a smaller value of the difference criterion corresponds to a better model of this TRC. A difference criterion introduces the relation of equivalence on a set of models of the certain fixed TRC and orders the equivalence classes linearly.

A difference criterion between actual TRC and its model must have two additional properties: symmetry property on its domain of definition and reflexive property ($trCr = mCr$ implies $DC(trCr, mCr) = 0$). These two properties are not imposed on the difference criterion between actual individual TRC and its model since semantics of the actual TRC values (including units of measurement) coincides with semantics of values of the TRC model, obtained by means of the VS-model, guaranteed for generalized TRCs only.

The necessary properties of a difference criterion do not guarantee its continuity. Multivalued difference criteria have not been considered in this article.

The testing and approbation of the parameterization algorithm for the VS-model, which is presented below, have been done with the following two difference criteria which are being offered for using as default difference criteria. Difference criterion DC_{ITRC} between actual individual TRC and its model, and difference criterion DC_{GTRC} between actual generalized TRC and its model are defined with formulas:

$$DC_{ITRC}(trCr, mCr) = 1 - \text{crln}(trCr, mCr) + 1,25(1 - \text{sync}(trCr, mCr))^2,$$

$$DC_{GTRC}(trCr, mCr) = DC_{ITRC}(trCr, mCr) + 0,4\left(\max_{i \in Yrs} |trCr_i - mCr_i|\right)^2,$$

where $\text{crln}(trCr, mCr)$ and $\text{sync}(trCr, mCr)$ are coefficients of Pearson's correlation and of synchronism (coincidence) between vectors $trCr$ and mCr , $trCr_i$ (mCr_i) is a corresponding to the year i component of the vector $trCr$ (mCr).

The selection of such difference criteria is conditioned by a desire: 1) to get calculated TRCs the most possible positive correlated with their respective actual TRCs;

2) of not having small values of synchronism coefficient between actual TRC and its model; 3) of absence of serious visual differences between broken lines that represent an actual generalized TRC and its model. The difference criteria DC_{TRC} and DC_{GTRC} have been tested on extensive data, are compatible with the parameterization algorithm for the VS-model (see below), and reflect the view of most researchers on concept of proximity of two TRCs.

The parameterization algorithm. Let us introduce definitions and denotations. Let $p = (p_1, \dots, p_n)$ denote vector of the VS-model parameters used under modeling (PUuM). Vector p is uniquely defined by a set of the VS-model options values used under modeling. We denote that the numbers of these n parameters of the VS-model, which certain numerical values are known for, by i_1, \dots, i_k ; let c_{i_1}, \dots, c_{i_k} denote the certain numerical values known for these parameters. The range of values of the VS-model parameter i is denoted by $[a_i; b_i]$. For the ease of exposition, it is supposed that $a_{i_j} = b_{i_j} = c_{i_j}$ for $j = 1, \dots, k$.

Let us call the following subset of $(n - k)$ a dimensional parallelepiped:

$$P = \{p \in \mathbf{R}^n : a_i \leq p_i \leq b_i \text{ for } 1 \leq i \leq n\}$$

lying in \mathbf{R}^n the optimization space S . S is defined by the condition: an element of P , considered as a set of values of the PUuM, belongs to S if and only if it has the necessary internal structure (see above). Call a set, belonging to S , of values of the PUuM a feasible. We consider $S \neq \emptyset$. The structure of S is conditioned by the interrelations, used under modeling, between the PUuM. S resembles a rectangular piece of cheese. It is a typical situation: among 100 000 values of the continuous n -dimensional random variate uniformly distributed on P only one belongs to S .

Define the predicate function $pFail(p): P \rightarrow \{\text{"TRUE"}, \text{"FALSE"}\}$ taking the logical value "TRUE" only if $p \notin S$. As a rule, for calculation of $pFail(p)$ the VS-model output data obtained for the set p are necessary.

Let us call the quantity:

$$\sup \{x \in \mathbb{R}^+ : \forall \text{InpOpt}_1 \forall \text{InpOpt}_2, \\ (|p_i^1 - p_i^2| \leq x \rightarrow Q(\text{Output}_1, \text{Output}_2))\}$$

the accuracy along the i -th axis, $1 \leq i \leq n$. Here InpOpt_1 and InpOpt_2 – are two collections of the VS-model input data and of values of its options, which differ only in values p_i^1 and p_i^2 of the i -th PUuM, Output_1 and Output_2 – the VS-model output data corresponding to InpOpt_1 and InpOpt_2 respectively. The predicate Q is true if and only if Output_1 and Output_2 differ from each other so little that this difference may be ignored and is considered to be negligible. Denote by h_i an estimate, used under modeling, of the accuracy along the i -th axis.

We introduce S – a metric $d: S \times S \rightarrow [0; +\infty)$ defined as:

$$d(x, y) = \max_{1 \leq i \leq n} \left| \frac{x_i - y_i}{h_i} \right|.$$

Selecting a metric on S is not a trivial problem. Such a metric is introduced in S since: 1) the optimization space has, generally speaking, different physical units of measurement along with different axes; 2) the constants h_i corresponding to axes with equal physical units of measurement can be different; 3) it is a natural analog of the metric ρ_∞ on \mathbf{R}^n [6].

Let us state the parameterization algorithm for the VS-model. The proposed algorithm selects a set of values of the PUuM from the optimization space S ; the selected set provides the global minimum over S of the objective function $F(p)$. The algorithm requires that $F(p)$ is a real single-valued nonnegative function defined everywhere on S . The algorithm does not require any additional properties of $F(p)$. The objective function is defined by a following equality:

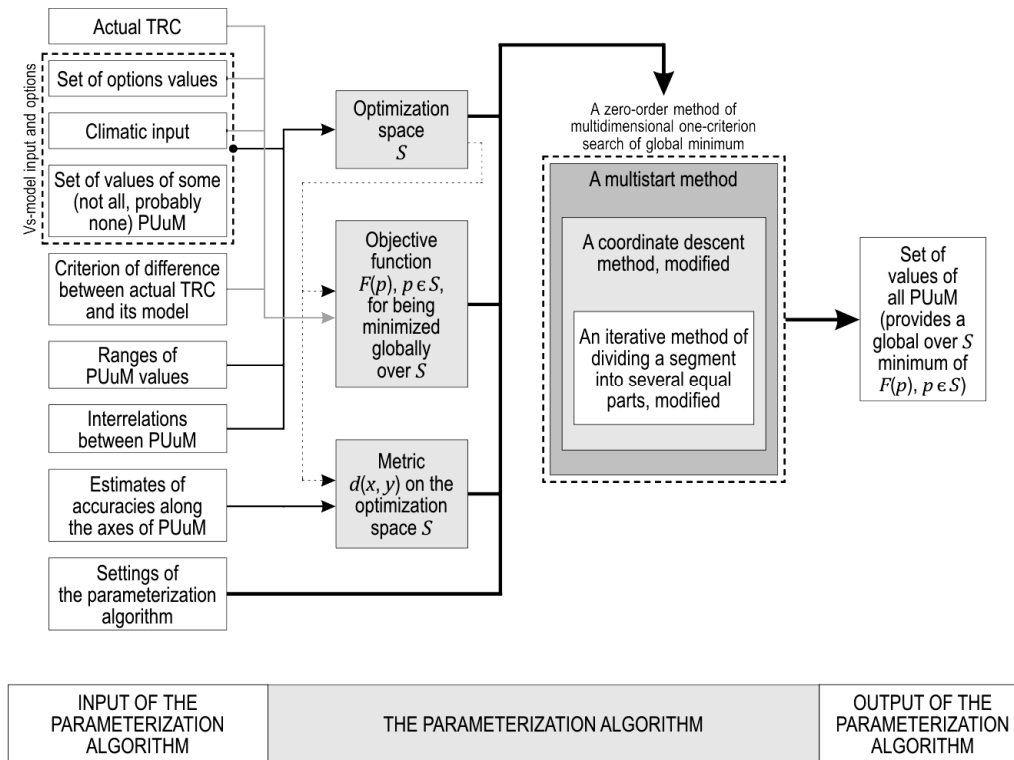
$$F(p) = DC(\text{trCr}, m\text{Cr}(p)), p \in S,$$

where $m\text{Cr}(p)$ is the vector of values, relating to the years Yrs , of the TRC model (see above) derived by using the VS-model with the set p of the PUuM values. The proposed parameterization algorithm for the VS-model is applicable with any difference criterion. It considers only those properties of the difference criterion being minimized by the one in which any difference criterion has necessarily.

The objective function $F(p)$ and its domain of definition S are individual for every situation which the parameterization algorithm being stated is applied in. Investigating the problem of searching the global over S minimum of $F(p)$ have shown that $F(p)$ has more than one local minimum as a rule. Input and output algorithm data and its structure are presented in figure.

A zero-order method of multidimensional one-criterion search of global minimum, used by the parameterization algorithm, is a multistart method of a coordinate descent method [7]. The question of existence of a more effective method for solving a family of problems, searching the global over S minimum of $F(p)$ is not considered in this article. The used coordinate descent method differs from the classical one in the following ways: 1) order of tracing coordinate axes of S is defined by a user; 2) a metric different from the Euclidean one is used on S ; 3) the used one-dimensional optimization method performs a search of a global optimum; 4) absence of requirement of $F(p)$ being continuous, and complicated structure of S require a modification of the classical coordinate descent method.

Settings of the parameterization algorithm are a collection of six constants m , $level$, $maxIter$, ε , q , $maxStart$, and two vectors τ and $wght$. The algorithm uses the same values of its settings over a period of all its operations being performed. The following settings of the parameterization algorithm are being offered for using as its default settings: $m = 3$, $level < \inf_{p \in S} F(p)$, $maxIter = 500$, $\varepsilon \leq (n - k)^{-1}$, $q = 1$, $maxStart = 1000$, $\tau_i \leq h_i \cdot \min(1, \varepsilon)$ ($1 \leq i \leq n$), $wght_j = 0$ for $j = i_1, \dots, i_k$ (default values of remaining $n - k$ positive components of the vector $wght$ are not noticed in the article). The oddness of m , and truth of inequality $\tau_i \leq h_i \cdot \varepsilon$, $1 \leq i \leq n$, are necessary for correct functioning of the algorithm.



Structural scheme of the parameterization algorithm for the VS-model

The multistart method launches the coordinate descent method from $maxStart$ starting points in S with the same settings. The least of $maxStart$ local minima found is an estimate of a global minimum and is delivered as a result of the multistart method run. Starting points are generated in a random manner and are different values of the continuous n -dimensional random variate uniformly distributed on P .

The order of tracing the coordinate axes of S with the coordinate descent method is defined by the weights $wght$. Each iteration of the latter consists in performing one-dimensional minimizations along the coordinate axes of S that have positive weights; halting condition is checked at the end of the iteration. One-dimensional minimizations along the axes that have maximum weight are carried out in the first instance. Axes with equal positive weights are traced in the order induced by their numbers – the smaller the number, the earlier the minimization is made.

The minimization along axis i , $1 \leq i \leq n$, means searching a global minimum of function $f(x) = F(\dots, p_{i-1}, x, p_{i+1}, \dots)$ of a real variable x on a set that lies inside the segment $[a_i; b_i]$ and which is specified by predicate function $xFail(x) = pFail(\dots, p_{i-1}, x, p_{i+1}, \dots)$. After having performed the search, a value x^* at which global minimum of $f(x)$ occurs is being written into the i -th coordinate of the current point p in S . In entering the first iteration, the current point p coincides with the starting one which has been selected by the multistart method. Coordinates of p with zero weights are not altered over a period of all operations being performed by the coordinate descent method.

The halting condition holds in case of truth of at least one of conditions: 1) number of iterations have reached a threshold $maxIter$; 2) $F(p) < level$, where p is a current point at the end of iteration (after executing all one-dimensional minimizations of the iteration); 3) $d(p^N, p^{N-i}) < \epsilon$ for all $i = 1, \dots, q$ (N – the current iteration number, p^j – current point in S at the end of iteration j , q and ϵ – positive constants); 4) $d(p^N, p^{N-1}) = 0$.

The one-dimensional minimizations along the coordinate axes of S are carried out by the method that is a modification of an iterative method of dividing segments into several equal parts [7]. During each iteration the method divides current segment into m equal parts and either makes one of them a current segment or completes the operation informing the user about the necessity to change the value m . The iterations stop when the value, on the current segment, of the function being minimized becomes less than the value $level$, or when the length of the current segment becomes less than the value δ defined before launching the method.

All one-dimensional minimizations along axis i , is carried out in the course of running the coordinate descent method, are performed with $\delta = \tau_i$ ($1 \leq i \leq n$).

The realization and approbation of the parameterization algorithm. The parameterization algorithm for the VS-model is realized in programming language C++, using the means of the Standard C++ Library. Its textual code satisfies the C++03 standard (ISO/IEC 14882:2003). The executable code of the realization for a certain operating system can be derived without changing the textual code of the realization and is arranged as a dll-library (if the objective platform is an

OS of the Windows family) or a text-file (if the objective platform is an OS of the Unix family).

The main factors influencing the duration of parameterization for a certain TRC by means of the realization, being considered, of the parameterization algorithm are: 1) the set of values of the VS-model options used under modeling; 2) the power of the set of the years which modeling is performed on; 3) the interrelations between PUuM used under modeling; 4) the value of the quantity $n - k$ and the lengths of the values ranges of those PUuM which specific numerical values are not known for; 5) the difference criterion used under parameterization; 6) the used settings of the parameterization algorithm; 7) hardware and software of a device carrying out the code of the parameterization algorithm realization. Varying any of these factors, one can get a significant change (of many times) of the duration of parameterization.

Using laptop Asus F3JP with dual core processor T5300 (maximum clock frequency of each core is 1.73 GHz), the parameterization of a TRC by using the realization of the parameterization algorithm usually lasts from several hours to several days.

The proposed parameterization algorithm for the VS-model was used at more than 700 different pairs ($S, F(p)$). Having performed these parameterizations as a consequence: 1) a number of statements about the VS-model, both the new and the previously formulated, are first confirmed by a computing experiment; 2) a hypothesis about the relationship between the models of individual TRCs and the model of the generalized TRC (all the models are derived by using the VS-model with the same set of values of its options) is for the first time formulated and confirmed by a computing experiment; 3) the VS-model has been first used in full as the quality tester of individual and generalized TRCs which are going to be used in the role of mediators carrying climatic

information; 4) several properties of the considered parameterization algorithm have been established. The successful approbation of the parameterization algorithm and its realization is due to these results.

The results obtained in this article are demanded mainly in dendroclimatology and dendrochronology, and in mathematical modeling of processes taking place in woody plants. The created parameterization algorithm and its realization: 1) significantly extend the range of situations in which using the VS-model is already realizable in practice; 2) allow the testing and analyzing of the model at a qualitatively new level; 3) allowing to solve in practice inverse problems of restoring certain growing conditions for woody plants via its existing TRC.

References

1. Hughes M. K., Swetnam T. W., Diaz H. F. Dendroclimatology: Progress and Prospects (Developments in Paleocological Research) // Berlin : Springer-Verlag, 2009.
2. Vaganov E. A., Hughes M. K., Shashkin A. V. Growth Dynamics of Conifer Tree Rings: Images of Past and Future Environments. Berlin : Springer-Verlag, 2006.
3. Vaganov E. A., Shashkin A. V. Growth and tree-ring structure of conifers. Novosibirsk : Nauka, 2000.
4. Cook E. R., Kairiukstis L. Methods of Dendrochronology: applications in the environmental sciences. Dordrecht : Kluwer Acad. Publ., 1990.
5. Papalambros P. Y., Wilde D. J. Principles of Optimal Design: Modeling and Computation. Cambridge : Cambridge University Press, 2000.
6. Kolmogorov A. N., Fomin S. V. Elements of the Theory of Functions and Functional Analysis. N. Y. : Dover Publications, 1999.
7. Horst R., Pardalos P. M. Handbook of global optimization. Dordrecht : Kluwer Academic Publishers, 1995. Vol. 1.

© Ivanovsky A. B., Shishov V. V., 2010

D. V. Kapulin, I. S. Korosteleva, D. A. Pupaeva
Siberian Federal University, Russia, Krasnoyarsk

CALCULATION METHOD FOR AUTOMATIC SECONDARY POWER SOURCES*

The article describes methods for calculating the parameters of secondary power sources (SPS): rectifying devices and voltage impulse stabilizers. Algorithms for calculating parameters of SPS power circuits have been developed.

Keywords: secondary power sources, algorithms for calculating the parameters of secondary power sources.

Secondary power sources (SPS) are an integral part of any radio set. Modern SPSs of electronic equipment have developed far beyond the class of simple radio-electronic devices, containing a small number of elements as it was 25–30 years ago. Today the secondary power sources are fairly complex devices which contain a large variety of functional

units, performing certain functional transformations of electrical energy improving its quality. All this demands constant increase of performance from the SPS, while the time requirement for the design of the SPSs is being reduced. Electronic circuits of SPSs are characterized by a presence of components with nonlinear characteristics.

*The article has been written due to the Young Scientists Support Program of the Siberian Federal University.