

This discussion paper is/has been under review for the journal Hydrology and Earth System Sciences (HESS). Please refer to the corresponding final paper in HESS if available.

A strategy for “constraint-based” parameter specification for environmental models

S. Gharari^{1,2}, M. Shafiei^{1,3}, M. Hrachowitz¹, F. Fenicia^{1,2}, H. V. Gupta⁴, and H. H. G. Savenije¹

¹Delft University of Technology, Faculty of Civil Engineering and Geosciences, Water Resources Section, Delft, the Netherlands

²Public Research Center–Gabriel Lippmann, Belvaux, Luxembourg

³Department of Water Engineering, Faculty of Agriculture, Ferdowsi University of Mashhad, Mashhad, Iran

⁴Department of Hydrology and Water Resources, The University of Arizona, Tucson AZ, USA

Received: 10 November 2013 – Accepted: 19 November 2013 – Published: 5 December 2013

Correspondence to: S. Gharari (s.gharari@tudelft.nl)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

Many environmental systems models, such as conceptual rainfall-runoff models, rely on model calibration for parameter identification. For this, an observed output time series (such as runoff) is needed, but frequently not available. Here, we explore another way to constrain the parameter values of semi-distributed conceptual models, based on two types of restrictions derived from prior (or expert) knowledge. The first, called “parameter constraints”, restrict the solution space based on realistic relationships that must hold between the different parameters of the model while the second, called “process constraints” require that additional realism relationships between the fluxes and state variables must be satisfied. Specifically, we propose a strategy for finding parameter sets that simultaneously satisfy all such constraints, based on stepwise sampling of the parameter space. Such parameter sets have the desirable property of being consistent with the modeler’s intuition of how the catchment functions, and can (if necessary) serve as prior information for further investigations by reducing the prior uncertainties associated with both calibration and prediction.

1 Introduction

Environmental systems models, such as conceptual rainfall-runoff (CRR) models, are abstract simplifications of real system behavior. Often, the parameters in such models cannot be specified through direct measurements of physical properties of the system. Further, even when a parameter is related to measurable quantities, its value in the model typically represents an integrated value over a much larger scale than the measurement scale. For this reason, such models typically rely upon calibration (tuning to match system input-output behavior for some historical data period) to ensure satisfactory predictive performance when applied to specific hydrological systems of interest (Wheater et al., 1993; Beven, 2001).

Constraint-based parameter specification

S. Gharari et al.

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)

[◀](#)

[▶](#)

[◀](#)

[▶](#)

[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)



In the case of CRR, parameter values are typically specified through a process of calibration that seeks to match the model runoff simulations to observed hydrographs. Expert knowledge is brought to bear implicitly, by the prior specification of parameter ranges that define the feasible parameter space. Recently, several studies have tested strategies that relate the parameter values of CRR models to catchment characteristics (Koren et al., 2000, 2003; Anderson et al., 2006; Yadav et al., 2007; Pokhrel et al., 2008, 2012; Kling and Gupta, 2009); the general picture that emerges from these studies is that exploiting expert knowledge (by imposing more rigorous constraints on the parameters) has the potential to result in more realistic models (Martinez and Gupta, 2011) and therefore reduced predictive uncertainty.

As a specific example, Pokhrel et al. (2008, 2012) linked the parameters of a spatially-distributed model to physical catchment characteristics via a set of regularization relationships, thereby converting the original high-dimensional parameter estimation problem to one of optimizing a reduced dimensional set of “super-parameters”, thereby dramatically simplifying the problem. Similarly, Merz and Blöschl (2004), Kling and Gupta (2009) and Yadav et al. (2007), amongst others, investigated explicit links between physical catchment characteristics and the parameters of a simple lumped conceptual model; they concluded, however, that such relationships are challenging to establish and may not often be possible given the available data. Other studies have used a comparison of catchment characteristics based on similarities between catchments to help constrain parameter values; for example, Zhang et al. (2008) imposed a set of three constraints to infer the runoff characteristics of catchments without calibration to observed hydrographs.

A promising approach that has recently been investigated is the use of parameter regionalization relationships to infer model parameter values. Kapangaziwiri et al. (2012) constrained the Pitman monthly rainfall runoff model (Hughes et al., 2006) based on a regionalization of runoff signatures. Perrin et al. (2008) proposed a method called *discrete parameterization* based on the use of parameter sets compiled a priori via calibration to other catchments. Their approach “abandon(s) the idea of searching for an

optimum parameter set in the continuous, n-dimensional parameter space” and instead “limit(s) the calibration process to a search within a finite collection (a library) of predefined parameter sets”. More recently, Samaniego et al. (2010) and Kumar et al. (2010, 2013) demonstrated that a multi-scale approach to parameter regionalization can provide consistent model performance for both gauged and ungauged catchments.

In a complementary direction, the use of multiple objective functions or multiple system responses for calibration (Gupta et al., 1998) has been shown to result in more realistic parameter sets that achieve improved simulations of system dynamics for the right reason (cf. Kirchner, 2006). The multi-objective approach seeks to identify parameter sets that simultaneously provide “optimal” performance for different aspects of system response (Gupta et al., 1998; Boyle et al., 2000, 2001). This can include constraining the model to reproduce multiple system fluxes and state variables such as evaporation, groundwater levels, tracer concentrations etc. (e.g. Gupta et al., 1999; Bastidas et al., 1999; Freer et al., 2002; Seibert and McDonnell, 2002; Khu and Madsen, 2005; Fenicia et al., 2008; Winsemius et al., 2008; Birkel et al., 2011; Hrachowitz et al., 2013; Seibert and McDonnell, 2013).

While the aforementioned studies have demonstrated that incorporation of expert and a priori knowledge can help improve the realism of models, no systematic strategy has been presented in the literature for constraining the model parameters to be consistent with the (sometimes) patchy understanding of a modeler regarding how the real system might work. Part of the difficulty in doing this is that expert knowledge may not always consist of explicitly quantifiable relationships between physical system characteristics and model parameters; rather, it may consist of conceptual understanding about consistency relationships that must exist between various model parameter or behavioral relationships (Hornberger and Spear, 1981) that must exist among model state variables and/or fluxes. For example, the geology of a given catchment may suggest that the catchment response during intense rainfall events is characterized by a slow responding groundwater component accompanied by fast responding Hortonian overland flow. In such a situation, any model results that imply that peak flows are

Constraint-based parameter specification

S. Gharari et al.

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)

[◀](#)

[▶](#)

[◀](#)

[▶](#)

[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)



composed of a strong groundwater response should be discarded or given lower importance. Such information acts as a constraint on the set of feasible model behaviors, and can thus help to limit the feasible extent of the model parameter space, resulting in reduced parameter and predictive uncertainty. As pointed out by Efstratiadis and Koutsoyiannis (2010) “It also offers a means to partially handle the huge uncertainty resulting from the complexity of model parameterizations in contrast to data scarcity, which is a global engineering problem that is getting increasingly severe. Actual research should provide more guidance on the effective combination of statistical and expert-based evaluation procedures.”

Here, we present a “constraint-based” strategy for constraining the feasible parameter space of a conceptual model, based on relational constraints inferred from expert knowledge regarding plausible catchment behavior. The approach is applicable to both lumped/semi-distributed and spatially distributed catchment models.

2 Constraints in models

Constraints on a model are of two main types, *a priori* constraints on model parameters (i.e. *parameter constraints*) and *a posteriori* constraints on model states and fluxes (i.e. *process constraints*; e.g., see Bulygina and Gupta, 2009, 2010, 2011). Parameter constraints are considered to be *a priori* because they can be imposed without actually running the model, while process constraints can only be imposed after a model is run with selected parameter sets.

2.1 Parameter constraints

Parameter constraints provide information regarding the relationships between parameters of the same process that correspond to different spatial components of a (semi-) distributed model. Such constraints can be expressed by equality or inequality constraints; for example:

Constraint-based parameter specification

S. Gharari et al.

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)

[◀](#)

[▶](#)

[◀](#)

[▶](#)

[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)



$$A_1 < A_2 \quad (1)$$

$$A_1 < A_2 + N \quad (2)$$

where N has the same unit as A_1 and A_2 .

$$^5 \quad A_1 B_1 < A_2 B_2 \quad (3)$$

As a simple illustration of this concept, the maximum interception capacity of a forested area ($I_{\max, \text{forest}}$) can typically be assumed to larger than the maximum interception capacity of a grassland area ($I_{\max, \text{grass}}$).

2.2 Process constraints

- ¹⁰ Process constraints provide comparative information regarding the fluxes (F) and/or states (S) of a model at each time step, or integrated over some specific time period. Examples of such constraints include:

$$\int_{t_2}^{t_1} F_1 dt < \int_{t_2}^{t_1} F_2 dt \quad (4)$$

$$\frac{\int_{t_2}^{t_1} F_1 dt}{\int_{t_2}^{t_1} F_2 dt} < G \quad (5)$$

$$^15 \quad \frac{S_{1,t_1} dt}{S_{2,t_1} dt} < G \quad (6)$$

where G is a dimensionless constant.

As an illustration, one can compare the transpiration fluxes from different spatial entities of a (semi-) distributed model. For example for two regions having similar soil

Constraint-based parameter specification

S. Gharari et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Constraint-based parameter specification

S. Gharari et al.

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[◀](#)[▶](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

type and aspect, the region with smaller normalized difference vegetation index (NDVI) is expected to transpire at a lower rate.

Note that in either case, parameter sets that satisfy the constraints are not conditional on information provided by *observations* (measurements) of the output response of the system (e.g. the runoff hydrograph), and these can therefore be determined without resorting to model calibration. However, parameter sets that satisfy all of the constraints can provide insights into how the real system can be expected to behave, assuming that it corresponds to the expert's perception of realistic (behavioral) system properties and dynamics.

Unfortunately, the use of available evolutionary algorithms to search for parameter sets that satisfy such constraints is complicated by the non-convex and potentially non-continuous parameter search space that results. In Sect. 3 we propose a stepwise search strategy that can be used to identify parameter sets that satisfy the full set of conditions imposed by expert knowledge.

3 Methodology and algorithm

The method is based on a simple stepwise search for parts of the parameter space that satisfy the set of constraints as discussed in the previous section. At each step, the algorithm tries to generate new parameter sets that satisfy the parameter constraints, while only violating the process-based constraints to an “acceptable” level. The process continues until such time that all of the generated parameter sets properly satisfy of imposed process-based constraints.

In the following description M refers to the total number of process-based constraints and $m \in \{1, \dots, M\}$ is an index indicating how many of the process-based constraints are satisfied by a given parameter set; for example, if a parameter set satisfies two process-based constraints then m will be equal to 2. The algorithm ultimately generates a set P of n number of behavioral parameter sets that satisfy all of the parameter

and process-based constraints (i.e. $M = m$ for all of members of P). Figure 2 presents a graphical illustration of these steps.

- Step 0: Begin with $C = 2$.
- Step 1: Generate N random samples (parameter sets) across the entire feasible parameter space using uniform prior distributions.
- Step 2: Evaluate the parameter constraints and identify samples that satisfy them.
- Step 3: Run the model is run for the samples identified in step 2, evaluate the M process-based constraints for each samples and assign a value of m to each parameter set corresponding to the number of process constraints satisfied.
- Step 4: Place the samples that satisfy C or more process-based constraints in set P , and the those that satisfy exactly $C - 1$ process-based constraints in set P' . Discard samples that satisfy $C - 2$ or fewer process-based constraints.
- Step 5: Use the members of sets P and P' to generate K new samples by applying each of the three Monte Carlo based rules below to generate $K/3$ of the samples, where θ_{new} is the newly generated sample. θ_P and $\theta_{P'}$ are samples selected randomly from sets P and P' respectively and α is a random value between 0 and 1. Figure 1 shows a graphical illustration of these rules.

$$\theta_{\text{new}} = \alpha\theta_P + (1 - \alpha)\theta_P \quad (7)$$

$$\theta_{\text{new}} = \alpha\theta_P + (1 - \alpha)\theta_{P'} \quad (8)$$

$$\theta_{\text{new}} = \alpha\theta_{P'} + (1 - \alpha)\theta_P \quad (9)$$

- Step 6: Discard all existing members of set P' (i.e. $P' = \emptyset$).
- Step 7: Increase C by one and return to step 2. Repeat this process is repeated until C becomes equal to the total number of process-based constraints (i.e. $C = M$).

Constraint-based parameter specification

S. Gharari et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

|◀

▶|

◀

▶|

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Constraint-based parameter specification

S. Gharari et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Note that any member of set P is within the space marked by members of set P' . Using members of P' to generate new parameter sets (step 5) helps to identify the boundary between the parameter space that satisfies exactly $C - 1$ (set P') and C or more (set P) constraints. The intention is to obtain a diverse parameter representation for set P by including the set P' (Fig. 1).

The final set P contains parameter sets that satisfy all of the parameter and process constraints. These parameter sets can be referred to as *constrained but un-calibrated*, as they are not constrained based on observed data about the target variables.

Note that the set P can also be used to constrain a search for “optimal” parameter sets within this space of *constrained but un-calibrated* parameter sets. This is easily done by evaluating them based on model performance in regard to a target variable (e.g. observed runoff). As such, the set P can be used as an initial sample for any evolutionary algorithm. In this case, any new parameter sets generated by the evolutionary algorithm would need to be checked for both parameter and process constraints and only retained if they satisfy the entire set of constraints.

Gharari et al. (2013) demonstrate how the proposed stepwise search algorithm can be used to specify parameters for a complex conceptual hydrological model applied to a mesoscale catchment.

4 Conclusions

One of the most challenging tasks in the development of complex conceptual hydrological models is the specification of realistic values for the parameter values. We have presented a strategy that enables incorporation of expert knowledge (i.e. the modeler’s perception of catchment behavior and characteristics) into the parameter specification process. Because the algorithm does not require observational data regarding the target system output (e.g. runoff) it can provide an efficient way to bridge the gap in the dialogue between modelers and experimentalists. Further, it can help to provide behaviorally superior parameter sets when used in conjunction with model calibration.

Acknowledgements. Shervan Gharari is funded during his PhD program by Fonds National de la Recherche (FNR) of Luxembourg with Aides la Formation-Recherche (AFR) project number of 1383201. Mojtaba Shafiei is partially funded by Iran's Ministry of Science, Research and Technology as exchange researcher at Delft University of Technology.

5 **References**

- Anderson, R. M., Koren, V. I., and Reed, S. M.: Using SSURGO data to improve Sacramento Model a priori parameter estimates, *J. Hydrol.*, 320, 103–116, doi:10.1016/j.jhydrol.2005.07.020, 2006. 14859
- Bastidas, L. A., Gupta, H. V., Sorooshian, S., Shuttleworth, W. J., and Yang, Z. L.: Sensitivity analysis of a land surface scheme using multicriteria methods, *J. Geophys. Res.-Atmos.*, 104, 19481–19490, doi:10.1029/1999JD900155, 1999. 14860
- Beven, K. J.: Rainfall-runoff modelling: the primer, Wiley, Chichester, 2001. 14858
- Birkel, C. D. T., Dunn, S., and Soulsby, C.: Using time domain and geographic source tracers to conceptualize streamflow generation processes in lumped rainfall-runoff models, *Water Resour. Res.*, 47, W02515, doi:10.1029/2010WR009547, 2011. 14860
- Boyle, D. P., Gupta, H. V., and Sorooshian, S.: Toward improved calibration of hydrologic models: combining the strengths of manual and automatic methods, *Water Resour. Res.*, 36, 3663–3674, doi:10.1029/2000WR900207, 2000. 14860
- Boyle, D. P., Gupta, H. V., Sorooshian, S., Koren, V., Zhang, Z., and Smith, M.: Toward improved streamflow forecasts: value of semidistributed modeling, *Water Resour. Res.*, 37, 2749–2759, doi:10.1029/2000WR000207, 2001. 14860
- Bulygina, N. and Gupta, H.: Estimating the uncertain mathematical structure of a water balance model via Bayesian data assimilation, *Water Resour. Res.*, 45, W00B13, doi:10.1029/2007WR006749, 2009. 14861
- Bulygina, N. and Gupta, H.: How Bayesian data assimilation can be used to estimate the mathematical structure of a model, *Stoch. Env. Res. Risk A.*, 24, 925–937, doi:10.1007/s00477-010-0387-y, 2010. 14861
- Bulygina, N. and Gupta, H.: Correcting the mathematical structure of a hydrological model via Bayesian data assimilation, *Water Resour. Res.*, 47, W05514, doi:10.1029/2010WR009614, 2011. 14861

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[◀](#)[▶](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

- Efstratiadis, A. and Koutsoyiannis, D.: One decade of multi-objective calibration approaches in hydrological modelling: a review, *Hydrolog. Sci. J.*, 55, 58–78, doi:10.1080/02626660903526292, 2010. 14861
- Fenicia, F., Savenije, H. H. G., Matgen, P., and Pfister, L.: Understanding catchment behavior through stepwise model concept improvement, *Water Resour. Res.*, 44, W01402, doi:10.1029/2006WR005563, 2008. 14860
- Freer, J., McDonnell, J. J., Beven, K. J., Peters, N. E., Burns, D. A., Hooper, R. P., Aulenbach, B., and Kendall, C.: The role of bedrock topography on subsurface storm flow, *Water Resour. Res.*, 38, 1269, doi:10.1029/2001WR000872, 2002. 14860
- Gharari, S., Hrachowitz, M., Fenicia, F., Gao, H., and Savenije, H. H. G.: Using expert knowledge to increase realism in environmental system models can dramatically reduce the need for calibration, *Hydrol. Earth Syst. Sci. Discuss.*, 10, 14801–14855, doi:10.5194/hessd-10-14801-2013, 2013. 14865
- Gupta, H. V., Sorooshian, S., and Yapo, P. O.: Toward improved calibration of hydrologic models: multiple and noncommensurable measures of information, *Water Resour. Res.*, 34, 751–763, doi:10.1029/97WR03495, 1998. 14860
- Gupta, H. V., Bastidas, L. A., Sorooshian, S., Shuttleworth, W. J., and Yang, Z. L.: Parameter estimation of a land surface scheme using multicriteria methods, *J. Geophys. Res.-Atmos.*, 104, 19491–19503, doi:10.1029/1999JD900154, 1999. 14860
- Hornberger, G. M. and Spear, R. C.: Approach to the preliminary analysis of environmental systems, *J. Environ. Manage.*, 12, 7–18, 1981. 14860
- Hrachowitz, M., Savenije, H., Bogaard, T. A., Tetzlaff, D., and Soulsby, C.: What can flux tracking teach us about water age distribution patterns and their temporal dynamics?, *Hydrol. Earth Syst. Sci.*, 17, 533–564, doi:10.5194/hess-17-533-2013, 2013. 14860
- Hughes, D. A., Andersson, L., Wilk, J., and Savenije, H. H.: Regional calibration of the Pitman model for the Okavango River, *J. Hydrol.*, 331, 30–42, doi:10.1016/j.jhydrol.2006.04.047, 2006. 14859
- Kapangaziwiri, E., Hughes, D., and Wagener, T.: Incorporating uncertainty in hydrological predictions for gauged and ungauged basins in southern Africa, *Hydrolog. Sci. J.*, 57, 1000–1019, doi:10.1080/02626667.2012.690881, 2012. 14859
- Khu, S. T. and Madsen, H.: Multiobjective calibration with Pareto preference ordering: an application to rainfall-runoff model calibration, *Water Resour. Res.*, 41, W03004, doi:10.1029/2004WR003041, 2005. 14860

Constraint-based parameter specification

S. Gharari et al.

[Title Page](#)

[Abstract](#)

[Introduction](#)

[Conclusions](#)

[References](#)

[Tables](#)

[Figures](#)

◀

▶

◀

▶

[Back](#)

[Close](#)

[Full Screen / Esc](#)

[Printer-friendly Version](#)

[Interactive Discussion](#)



Constraint-based parameter specification

S. Gharari et al.

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[|◀](#)[▶|](#)[◀](#)[▶|](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

Constraint-based parameter specification

S. Gharari et al.

Title Page

Abstract

Introduction

Conclusion

References

Tables

Figures

Page 1

1

10 of 10

1

Back

Close

Full Screen / Esc

[Printer-friendly Version](#)

Interactive Discussion



Constraint-based parameter specification

S. Gharari et al.

Title Page

Abstract

Introduction

Conclusion

References

Tables

Figures

14

7

Back

Close

Fu

/ Esc

Interactive Discussion

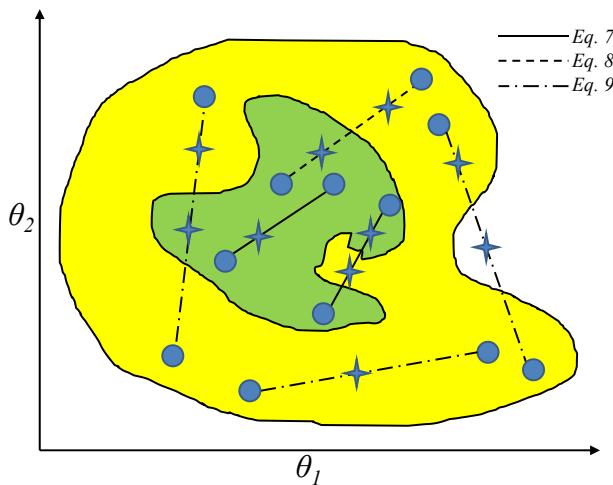


Fig. 1. A conceptual illustration of possible positions of newly generated parameter sets based on parameter sets randomly drawn from P and P' for a two dimensional parameter space. The area indicated by yellow represents the set P' that satisfies exactly $C - 1$ process based constraints. The area indicated by green represents the set of P that satisfies C or more process constraints. The circles indicate randomly selected parameter sets drawn from the sets P or P' . Different line style indicate different parameter generation rules (insert the appropriate equation numbers to the different lines). Solid lines represent the first rule where the parameter sets are randomly selected from set P (Eq. 7), dashed lines show the second rule where one parameter set is randomly selected from P and one from P' (Eq. 8), the dash-dot line represents the third rule where both randomly selected parameter sets are selected from set P' (Eq. 9). Note that due to possible non-convexity of sets of P and P' the newly generated parameter sets based on the three rules can be outside of sets P and P' .

Constraint-based parameter specification

S. Gharari et al.

Title Page	
Abstract	Introduction
Conclusions	References
Tables	Figures
	
	
Back	Close
Full Screen / Esc	
Printer-friendly Version	
Interactive Discussion	

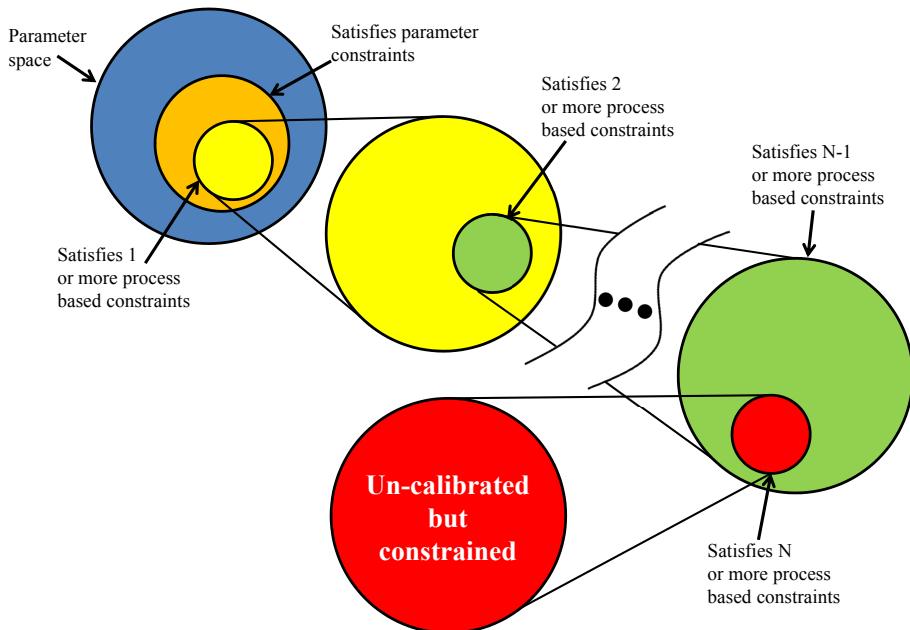


Fig. 2. A conceptual illustration of stepwise search for the parameter space satisfying all of the parameter and process constraints.