Stochastic Environmental Research and Risk Assessment 17 (2003) 276–290 © Springer-Verlag 2003
DOI 10.1007/s00477-003-0150-8

Another look at the conceptual fundamentals of porous media upscaling

G. Christakos

Abstract. We suggest a critical look at the epistemic foundations of the porous media upscaling problem that focuses on conceptual processes at work and not merely on form manipulations. We explore the way in which critical aspects of scientific methodology make their appearance in the upscaling context, thus generating useful effective parameters in practice. The fons et origo of our approach is a conceptual blending of knowledge states that requires the revision of the traditional method of scientific argument underlying most upscaling techniques. By contrast to previous techniques, the scientific reasoning of the proposed upscaling approach is based on a stochastic model that involves teleologic solutions and stochastic logic integration principles. The syllogistic form of the approach has important advantages over the traditional reasoning scheme of porous media upscaling, such as: it allows the rigorous derivation of the joint probability distributions of hydraulic gradients and conductivities across space; it imposes no restriction on the functional form of the effective parameters or the shape of the probability laws governing the random media (non-Gaussian distributions, multiple-point statistics and non-linear models are automatically incorporated); it relies on sound methodological principles rather than being ad hoc; and it offers the rational means for integrating the multifarious core knowledge bases and uncertain site-specific information sources about the subsurface system. Previous upscaling results are derived as special cases of the proposed upscaling approach under limited conditions of porous media flow, a fact that further demonstrates the generalization power of the approach. Our hope is that looking at the upscaling problem in this novel way will direct further attention to the methodological exploration of the problem at the length and the detail that it deserves.

Keywords: Upscaling, Porous media, Conceptual, Teleology, Logic, Uncertainty, Epistemic, Stochastic

G. Christakos
Center for the Advanced Study of the Environment,
Department of Environmental Sciences and Engineering,
School of Public Health, University of North Carolina,
Chapel Hill, NC 27599-7400, USA
e-mail: george_christakos@unc.edu

I would like to thank Drs. A. Kolovos and D.T. Hristopulos for their valuable comments. The work was supported by grants from the Army Research Office (Grant no. DAAG55–98–1-0289) and the National Institute of Environmental Health Sciences (P42-ES05948 & P30-ES10126).

1 Introduction

A commonly encountered hierarchy in physical sciences is that of natural scales, from the microscopic scale of particles and elementary volumes to the macroscopic scale of everyday life. Theoretical laws may be assumed at a microscopic level, whereas phenomenological laws are implemented at the control cell level of a geographical information system. Viewed from this angle, many physical entities are not absolute but rather obtain their meaning within the context of the specified hierarchical level. Such a multi-scale consideration of Nature has far-reaching implications for our knowledge of it and, thus, a strong epistemic component.

Porous media investigations often involve a change-of-scale technique which relates a physical parameter at different scales. Flow and transport in heterogeneous porous media, e.g., are represented by a set of primitive equations relating subsurface parameters at a small spatial scale. However, for various reasons (acquisition of reliable knowledge through experimentation or numerical simulation) flow and transport dynamics need to be considered at a large scale using effective parameters, like effective hydraulic conductivity (EHC) or equivalent hydraulic conductivity (EqHC). The relation between small scale measurements and effective parameters is assessed quantitatively by means of an upscaling procedure.

In hydrogeologic research and development we should distinguish between, (i) investigation techniques (such as solving a physical equation, constructing a simulation model, or designing an experimental procedure), and (ii) conceptual reasoning frameworks (such as developing a methodology for applying the laws of logic to hydrologic situations, building hypotheses, or integrating physical knolwedge bases). As regards (i) above, in the upscaling literature we find various upscaling techniques often based on different interpretations of upscaled conductivity (e.g., Cushman et al., 2002). Two of the most common groups of such techniques are as follows:

a. The group of analytical techniques leading to the EHC interpretation. This group seeks to express EHC as function of conductivity statistics and most of its techniques are of a stochastic nature (Gelhar and Axness, 1983; Cushman, 1986; Dagan, 1989; Neuman et al., 1992; Neuman and Orr, 1993; Christakos et al., 1993, 1995; Hristopulos and Christakos, 1997a, b, 1999).
b. The group of numerical techniques leading to the EqHC interpretation of an upscaled conductivity. The flow equations apply in a controlled and self-contained part of the porous media whole in isolation of the rest. In this case, one often distinguishes between two sub-groups of techniques below:
   b1. one sub-group generates solutions of the primitive flow equation at the fine scale in terms of flux and hydraulic gradient distributions, and then derives the EqHC at the coarse scale in terms of some rule involving these distributions (e.g., Rubin and Gomez-Hernandez, 1990; Kitanidis, 1990; Durlofsky, 1991); and
   b2. another sub-group uses empirical power-averaging to calculate EqHC at the coarse scale without any explicit consideration of the flow conditions (e.g., Deutsch, 1989; Desbarat, 1992; Scheibe and Yabusaki, 1998).

While these techniques have offered valuable insight into certain aspects of the upscaling problem, their application in practice has been met with rather limited success for a number of reasons listed in the relevant literatue, including restrictive assumptions (like homogeneity), infinite flow domains and uniform hydraulic gradient, and low-order perturbation approximations.

As a matter of fact, all the investigations of the upscaling problem (groups a and b above) have focused mostly on form manipulations (analytical and computational aspects of the problem) and no much attention has been given to the underlying conceptual issues (e.g., inferences drawn in accordance with laws of logic, integration of upscaling within a broader framework of assumptions about the subsurface environment, scales that are connected in a phenomenological manner vs. a conceptual manner, and physical knowledge blending). Methodologically induced conceptual issues, however, can raise strong doubts about the problem solving efficacy of a wide range of formal techniques. Ergo, it is appropriate to revisit the fundamentals of the upscaling problem, viz., the critical reasoning concepts and methodological credentials underlying the upscaling techniques. This is a goal of the present work.

At this point, we should recall that the effective parameters obtained by the two major groups of techniques above are based on the traditional reasoning scheme of hydrologic science. In accordance with this scheme, one generally starts with a limited set of conductivity measurements at the fine scale and the models of the relevant low-order statistics (spatial mean and covariance or variogram) are generated. Then, by means of induction global validity of the statistics models is assumed, in which case two possible paths of action emerge:

A. The models are used in the context of a stochastically formulated equation of flow and specific solutions (usually approximate) for the EHC are sought by conventional means (see, also, group a above). In some cases, solutions in terms of statistical moments are conditioned to hard measurements at selected points in space.
B. The models are used to generate conductivity realizations at the fine scale by means of, e.g., geostatistical kriging. The realizations are then inserted:
   (B1) either in the flow equations that are solved at the fine scale to obtain the flux and gradient distributions over the domain, which in turn are used to obtain estimates of the EqHC by means of an upscaling rule (sub-group b1 above);
   (B2) or in a power-averaging rule to produce EqHC estimates (sub-group b2).

We would like to cast doubt on any idea that we can take the preceding upscaling techniques as givens that are in no way in need of methodological interpretation or critique. Instead, as already mentioned most of these techniques have significant limitations. It is worth-noticing, e.g., that no reliable procedure has been developed on the basis of the traditional reasoning scheme above to compute the joint probability distributions of hydraulic gradients and conductivity fields (SP Neuman Personal Communication). In this work, a different methodological scheme is considered which can eliminate such limitations. We focus on an epistemic treatment of upscaling, which is a return to the problem basics that can considerably expand the possible solutions. This treatment acknowledges the fact that the upscaling affair is not merely an assemblage of data and formal techniques. Instead, there is a logic to thinking in the upscaling context, which leads us to study the nature of the frame of thought by means of a number of epistemic concepts (blending, uncertainty, teleologic, information etc.) and their relations. A theoretical framework of upscaling is introduced which involves the integration of different input states such as core knowledge, incompletely known input parameters, site-specific uncertain influences and interdependencies to create new emergent structures, techniques, and ways of thinking. This is a knowledge synthesis approach which has a metacognitive character (it involves thinking about our thinking) and recognizes that behind form manipulation is human power to construct meaning by conceptual blending. By properly organizing physical concepts and logical processes underlying the porous media situation, knowledge synthesis can avoid limitations of previous upscaling techniques based on restrictive flow conditions, small heterogeneity assumptions, bounded domains, low-order perturbations, inadequate flow simulators or transfer functions, arbitrary upscaling transformations, and neural networks with unrealistic training datasets.

2 Stochastic formulation of knowledge bases

Due to uncertainty, the physical laws and mechanisms underlying a realistic hydrogeologic situation, even when they are known, manifest themselves in a complex manner which can be described only in stochastic terms. Stochastically formulated geologic media properties and physical laws account for the uncertainties generated when they are considered in a real-world environment and provide the range of possible values together with the probabilities of their occurrence across space. Stochastic formulation of the upscaling situation starts with the representation of the hydrologic parameters involved in terms of random fields across space and proceeds with the definition of a knowledge base (KB) as a collection of uncertain information sources relevant to the problem at hand which are invoked by a reasoning process aimed at solving the problem. Then, an efficient classification of KB can be established as follows.

2.1 General KB

The general KB $\mathcal{G}$ includes core knowledge and theoretical models developed for well-defined conceptual environments (fundamental laws of flow, primitive equations, etc.). Thus, $\mathcal{G}$-KB is often associated with science seeking to deepen insight at a fundamental level, in which case the physical laws of the KB give to the upscaling approach a nomological character (in the sense of Hempel).

For the purposes of the present study, the $\mathcal{G}$-KB includes the following. Consider a situation of effective flow in a porous domain which is sufficiently characterized by the local mean value law $(i = 1, \ldots, n;\ \alpha = 1, \ldots, N)$

$$\overline{K_\alpha J_{\alpha,i}} = K_{\mathrm{eff},\alpha,i} \overline{J_{\alpha,i}} \tag{1}$$

in some coordinate system $\mathbf{s} = (s_1, \ldots, s_n)$, where the bar denotes stochastic expectation, K_α is the random conductivity field $K(\mathbf{s})$ at point $\mathbf{s}_\alpha$, $\overline{J_{\alpha,i}}$, is the mean hydraulic gradient in the i direction expressed in terms of boundary conditions (BC) and conductivity statistics, and the $K_{\mathrm{eff},\alpha,i}$ are the EHC components sought. Depending on the situation, Eq. (1) can be associated with different BC. Equation (1) is a local law. Nonlocal laws may be considered as well (e.g., for nonhomogeneous situations), but this is beyond the methodological scope of the present work.

In addition to the physical Eq. (1), in most hydrogeologic situations the $\mathcal{G}$-KB can include theoretical models for 1-, 2- and multiple-point conductivity statistics of the general form

$$\beta_{\boldsymbol{\alpha}}^{(\boldsymbol{\rho})} = \overline{K_{\alpha_1}^{\rho_1} K_{\alpha_2}^{\rho_2} \cdots K_{\alpha_q}^{\rho_q}}\ , \tag{2}$$

where $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_q)$, $\alpha_1, \ldots, \alpha_q = 1, \ldots, N$, $\boldsymbol{\rho} = (\rho_1, \ldots, \rho_q)$ and $\rho_1, \ldots, \rho_q \in R$. Multiple-point nonlinear statistics provide information about the spatial variation of hydraulic conductivity that cannot be obtained from the 1- and 2-point statistics assumed by the conventional upscaling techniques. The $K_{\mathrm{eff},\boldsymbol{\alpha},i}$ field in Eq. (1) refers to one (large) scale, whereas the K_{α} field in Eq. (2) to another (small) scale. In the following, an issue of considerable methodological interest will be how knowledge contained in Eqs. (1) and (2) can be integrated within a broader framework of assumptions and models about the subsurface environment.

2.2
Specificatory KB

This KB includes site-specific details of the real subsurface environment in the form of hard data, secondary information sources and uncertain measurements (Christakos et al., 2002). In a sense, $\mathscr{S}$-KB states factual conditions (facts that make the laws of the $\mathscr{G}$-KB apply etc.). The $\mathscr{S}$-KB may include: an updated assessment of the site-specific conductivity domain, I_{κ}, at each spatial location; and/or the associated probability density function (pdf) of the conductivity field at some points across space, i.e.,

$$f_{\mathscr{S}}(\boldsymbol{\kappa}) = P\mathrm{d}f[K(\mathbf{s})] \ , \tag{3}$$

and/or an estimate of the pdf of the hydraulic gradient field at a set of points across space, i.e.,

$$f_{\mathscr{S}}(\boldsymbol{\zeta}) = P\mathrm{d}f[J(\mathbf{s})] \ , \tag{4}$$

where $\boldsymbol{\kappa}$ and $\boldsymbol{\zeta}$ are conductivity and hydraulic gradient realizations, respectively. The pdf (3), (4) can have any shape and domains $I_{\boldsymbol{\kappa}}$, $I_{\boldsymbol{\zeta}}$. The pdf are constructed on the basis of the ad rem conductivity measurements, physical intuition, etc. Other types of site-specific data bases are possible, including secondary information in the form of geophysical and geologic data (Serre and Christakos, 1999). The available hydraulic conductivity measurements could be of different scales (core scale, slug and packer tests, short- and long-term pumping tests, etc.). The $\mathscr{S}$-KB usually refers to science as a basis of action (aiming at predictive prediction of open systems, etc.). When the physical laws in $\mathscr{G}$-KB are considered in the context of an open system, to have practical meaning certain aspects have to be clarified, like the domain of application of the laws, the error boundaries within which the laws predictions are acceptable, and consistency with the data in $\mathscr{S}$-KB.

3
Upscaling as knowledge synthesis

Efficient EHC computations rely on the development of a sound methodology of upscaling. As was mentioned in Sect. 1 above, our approach to the upscaling problem involves an epistemic framework (i.e., it refers to processes by which knowledge and understanding are achieved, communicated and integrated within the scientific reasoning context). In this framework a distinction should be made between the term “method” and the term “methodology.” While the former term more commonly implies an orderly, step-by-step, prescriptive process with a predictable outcome, the latter term is indicative of a process that is based more on a set of general guiding principles, concepts and reasoning rules than a series of steps. In upscaling situations an adequate methodology can provide answers to important questions like "How can one make rigorous inference in view of general and site-specific KB?"

3.1
The basic methodological framework

In a real world study one often seeks the rational means to integrate core scientific knowledge, background experience and site-specific uncertain data into a unique methodological framework for upscaling purposes. Knowledge synthesis (KS) refers to the conceptual blending and processing of physical KB to obtain a realistic representation of the phenomenon across space, assess important uncertainty sources, evaluate relevant risks, and make science-based decisions. The credenda of the KS framework of porous media upscaling proposed in this work includes the integration of the $\mathcal{G}$-KB (abstract representation, pure thought and critical argument) with the $\mathcal{S}$-KB (appearance, limited experience and common sense) in order to generate a realistic and informative probability model linking the conductivity and flow gradient fields across space.

The KS reasoning assumes four stages of knowledge acquisition, blending, and processing, as shown in Table 1 (Christakos, 2000, 2002a). A comparison of the KS approach with the traditional reasoning scheme discussed in section 1 ante reveals certain important differences between the two. The KS methodology involves a stochastic logico-nomological model, in which objectivity refers to an ability to accept statements about the subsurface environment on the basis of scientific reasoning and empirical evidence. To improve epistemic objectivity and create a more realistic representation of the situation the KS methodology requires that one "looks" at the upscaling situation in two different ways: one with the help of the $\mathcal{G}$-KB, thus creating a first picture of the situation (model $f_{\mathcal{G}}$); and another one with the help of the $\mathcal{S}$-KB, thus creating a second picture of the situation (e.g., in terms of $f_{\mathcal{S}}$). Both pictures are necessary for the full understanding of the upscaling situation. Model $f_{\mathcal{G}}$ is obtained by means of an epistemically motivated teleologic rule (i.e., a purpose-oriented criterion involving an action principle) rather than by a direct solution of the stochastic flow equations by conventional means. While in most well-known action principles (including Aristotle's principle of minimum potential energy, Fermat's principle of least time, and Hamilton's principle of stationary principal function) the action

Table 1. The four stages of KS-based upscaling

Stochastic formulation of general knowledge: The general KB representing structural aspects of the subsurface environment is formulated stochastically in terms of a set of so-called $\mathcal{G}$-equations.
Model generation: A teleologic solution of the $\mathcal{G}$-equations is sought by means of an epistemic action principle (maximum expected information), thus leading to the pdf model, $f_{\mathcal{G}}$, of conductivity and flow gradients across space.
Site-specific knowledge formulation: The specificatory KB is transformed into a set of $\mathcal{S}$-operators representing the manifold case-specific aspects of the subsurface environment.
Revised model generation: The results of the previous stages are integrated to yield an updated pdf model, $f_{\mathcal{K}}$, of the conductivity and flow gradient fields ($\mathcal{K} = \mathcal{G} \cup \mathcal{S}$ denotes total KB). The EHC can be then expressed in terms of the porous media flow statistics calculated from the model $f_{\mathcal{K}}$.

sought refers to ontologic concepts (like energy and time), in the teleologic stage of Table 1 the action refers to an epistemic concept (information) that may be expressed mathematically in terms of the Shannon, the Fisher etc. information measures across space (Christakos et al., 2002). The two previous pictures are then blended by means of a *logic integration* rule to generate the final picture (model $f_{\mathcal{K}}$). The KS organon is very general allowing the use of different systems to yield an updated pdf model $f_{\mathcal{K}}$, including statistical inductive and stochastic deductive inferences (the choice of a system over another depends on the particular features of the porous media application considered). In a sense, the whole point of the methodological rules (i.e., teleologic and integration) is to offer norms for scientific behavior; to provide guidance in order to achieve the goal of the study, namely, the derivation of a richer theory of upscaling than previous techniques have allowed. In the following we examine these rules in some detail.

3.2 Teleologic solution and logic integration system

As was mentioned above, the teleologic pdf model $f_{\mathcal{G}}$ of the joint conductivity $\boldsymbol{\kappa}$ and hydraulic gradient $\boldsymbol{\zeta}$ distributions across space is derived by means of the epistemologically motivated action principle of maximum informativeness. In this novel context, whereas before the goal was to solve the flow equations in the conventional sense, this problem has no longer, strictly speaking, any significance. If the Shannon information concept is used in terms of the random fields across space, the teleologic model has the functional form (mathematical derivations are presented in Christakos, 2000)

$$f_{\mathcal{G}}(\boldsymbol{\kappa}, \boldsymbol{\zeta}) = e^{\boldsymbol{\mu}^{T}\mathbf{g}} \ , \tag{5}$$

where $\mathbf{g} = \{g_\alpha; \alpha = 0, 1, \ldots, N\}$ is a vector of g_α-functions of the gradient and conductivity fields properly chosen to express the physical $\mathcal{G}$-KB considered. E.g., as is shown in Appendix A the g_α-functions corresponding to Eqs. (A.1) are expressed by $\kappa_\alpha\zeta_{\alpha,i}$, $\zeta^{\lambda}_{\alpha,i}$ and $\Pi^{q}_{i=1}\kappa^{\rho_i}_{\alpha_i}$. The vector $\boldsymbol{\mu} = \{\mu_\alpha\}$ consists of coefficients associated with $\mathbf{g}$ (the coefficients $\mu_{\boldsymbol{\alpha}}$ are unknown, at this point, but are computed at a later KS stage, as is discussed in Sect. 3.3 below and in Appendix A). Generally, the $\mu_{\boldsymbol{\alpha}}$ are functions of the spatial coordinates and their initial values depend on the BC associated with the flow law (e.g., Kolovos et al., 2002).

Subsequent integration of site-specific information into upscaling yields an update of the teleologic model (5) by means of either of two groups of approaches (Christakos, 2000, 2002a): (i) operational Bayesian conditionalization (bc); and (ii) deductive inference. Group (i) is based on inductively strong standards, whereas group (ii) relies on deductively sound principles (the basic mathematical theory of deductive random fields is presented in Christakos, 2002b). The choice of an adequate integration approach is primarily a conceptual modelling affair supported by the physical and logical features of the situation. Operational bc is a powerful and versatile approach that uses knowledge-based probability operators to improve the teleologic model (5), thus leading to the following general model of the joint pdf of conductivity and hydraulic gradient across space,

$$f^{\mathrm{bc}}_{\mathcal{K}}(\boldsymbol{\kappa}, \boldsymbol{\zeta}) = \Theta_{\mathcal{S}} f_{\mathcal{G}}(\boldsymbol{\kappa}, \boldsymbol{\zeta}) \ , \tag{6}$$

where the operator $\Theta_{\mathcal{S}}$ depends on the kind of the $\mathcal{S}$-KB used. If, e.g., the $\mathcal{S}$-KB involves the site-specific probabilities (3) across space, then $\Theta_{\mathcal{S}} = A^{-1} f_{\mathcal{S}}(\boldsymbol{\kappa})$, where A is a normalization coefficient. Equation (6) incorporates information about the flow variables into the upscaling procedure without the need to solve any inverse problem.

In some cases, a deductive logic-based conditionalization may provide a better description of the relationship suggested by the laws of nature than a bc based on explicitly statistical reasoning or inductively strong standards. Deductive random field theory considers various shades of such a relationship across space and yields efficient rules that establish causal relevance in a physical sense. A useful deductive rule is material biconditionalization (mb), in which case the corresponding mb model of the joint pdf, $f_{\mathscr{K}}^{\mathrm{mb}}$, of conductivity and hydraulic gradient fields is (Christakos, 2002a, b)

$$f_{\mathscr{K}}^{\mathrm{mb}}(\boldsymbol{\kappa}, \boldsymbol{\zeta}) = \frac{1}{2A-1}[2A\Theta_{\mathscr{S}} - 1] f_{\mathscr{G}}(\boldsymbol{\kappa}, \boldsymbol{\zeta}) \ . \tag{7}$$

Probability models of the form of Eqs. (6) and (7) are generated at each point, thus providing a stochastically complete characterization of the upscaling situation across space rather than a single realization. Part of the methodological framework leading to these models consisted in treating the logic of upscaling as a theory of the operation of the faculty of reason, a faculty that acted to synthesize concepts and knowledge bases.

3.3
Operational expressions of effective hydraulic conductivity

The methodology briefly discussed above can have a decisive influence on the formulation of the upscaling problem itself, as well as on the nature of any possible solution to it. In order to obtain useful mathematical EHC formulae we need to continue expressing methodological arguments in terms of equations. Accordingly, after the probability model ($f_{\mathscr{K}} = f_{\mathscr{K}}^{\mathrm{bc}}$ or $= f_{\mathscr{K}}^{\mathrm{mb}}$) of the joint distribution of conductivity and flow gradient across space has been derived from Eqs. (6) and (7), the relevant porous media flow statistics can be calculated and used to obtain an expression for EHC. Hence, the next important step is the calculation of the vector $\boldsymbol{\mu}$. In light of the KS methodology (Sects. 3.1, 3.2) and given the $\mathscr{G}$- and $\mathscr{S}$-KB (Sect. 2), we define the operators

$$\left.\begin{aligned} \Lambda_{\mathscr{K}}[\cdot] &= \int_{I_{\boldsymbol{\kappa}}} \mathrm{d}\boldsymbol{\kappa} \int_{I_{\zeta}} \mathrm{d}\boldsymbol{\zeta}\, \Theta_{\mathscr{K}} \mathrm{e}^{\boldsymbol{\mu}^{\mathrm{T}}\mathbf{g}}[\cdot] \\ \Delta_{\mathscr{K}}[\cdot] &= \left(\int \mathrm{d}\boldsymbol{\kappa} \int \mathrm{d}\boldsymbol{\zeta} - \int_{I_{\boldsymbol{\kappa}}} \mathrm{d}\boldsymbol{\kappa} \int_{I_{\zeta}} \mathrm{d}\boldsymbol{\zeta}\, \Theta_{\mathscr{K}}\right) \mathrm{e}^{\boldsymbol{\mu}^{\mathrm{T}}\mathbf{g}}[\cdot] \end{aligned}\right\} \ . \tag{8}$$

As is shown in Appendix A, the vector $\boldsymbol{\mu}$ is the solution of the system of equations

$$\left.\begin{aligned} &\Delta_{\mathscr{K}}[\kappa_{\alpha}\zeta_{\alpha,i}] = 0, \qquad \Delta_{\mathscr{K}}[\zeta_{\alpha,i}^{\lambda}] = 0 \\ &\Delta_{\mathscr{K}}[\kappa_{\alpha_1}^{\rho_1}\kappa_{\alpha_2}^{\rho_2}\cdots\kappa_{\alpha_q}^{\rho_q}] = 0, \quad \Delta_{\mathscr{K}}[1] = 0, \quad \Lambda_{\mathscr{K}}[1] = 1 \end{aligned}\right\} \tag{9}$$

for all i, $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_q)$, $\lambda = 1, \ldots, L$, and $\boldsymbol{\rho} = (\rho_1, \ldots, \rho_q)$; the parameter $\Theta_{\mathscr{K}}$ is given by

$$\Theta_{\mathscr{K}} = \begin{cases} \Theta_{\mathscr{S}}, & \text{if bc is assumed} \\ \frac{1}{2A-1}[2A\Theta_{\mathscr{S}} - 1], & \text{if mb is assumed} \end{cases} \ . \tag{10}$$

Finally, the EHC expression is given in terms of operators (8) as follows (Appendix A)

$$K_{\mathrm{eff},\alpha,i} = \Lambda_{\mathscr{K}}[\kappa_\alpha \zeta_{\alpha,i}]\Lambda_{\mathscr{K}}^{-1}[\zeta_{\alpha,i}] \quad . \tag{11}$$

From a methodological viewpoint, Eq. (11) provides a conceptually revised definition of the EHC as a quantity that satisfies the teleologic and integration rules of the KS framework. Ergo, Eq. (11) is based on the integration of general laws and multiple-point statistics with uncertain site-specific data. The KS framework leading to Eq. (11) can rigorously account for hydraulic conductivity and flow gradient statistics across space in an implicit manner, which does not require the explicit involvement of these statistics. On the other hand, the choice of the initial $\boldsymbol{\mu}$-values in the numerical solution of Eqs. (9) is optimized by using information about the flow BC and conductivity statistics.

In a rather laconic manner, Table 2 presents the step-by-step KS treatment of the upscaling problem in practice. KS leads to a very general upscaling approach which, in principle, can produce EHC values for a congeries of flow conditions (spatially non-homogeneous fields, bounded domains, strongly heterogeneous media, etc.) and uncertain site-specific information sources (soft conductivity measurements, flow data, etc.). This approach has significant advantages over previous upscaling technique. A well-known group of stochastic techniques relies on perturbation approximations of the $\overline{K_\alpha J_{\alpha,i}}$ terms involving conductivity statistics. The KS approach follows a different path that formulates the upscaling problem in a way that there is no need for perturbation approximations, which is another significant advantage of the KS reasoning. Moreover, some stochastic hydrology techniques combine the averaged flow model with hard measurements at selected points through forward and/or inverse conditioning (FIC) of the former on the latter, coupled with what is often characterized as a "good measure of subjective reasoning" – although it is not always clear what this characterization implies. The methodological standards underlying the FIC approach are fundamentally different than those of the KS approach. KS relies on conceptual blending principles (seeking complete probability models with high information content to express a wide range of core knowledge bases and critical reasoning tools to assimilate the manifold site-specific data sources that extend well beyond hard measurements) rather than on the form manipulations of FIC (limited to low-order moment conditionalization by approximate spatial estimation techniques such as kriging, etc.). Furthermore, KS provides solutions to methodological issues of considerable interest such as, e.g., how can Eq. (1) be integrated within the broader framework of models and assumptions about the subsurface environment ($\mathscr{G}$- and $\mathscr{S}$-KB, etc.).

3.4
Some numerical illustrations

Any combination of core and application-specific KB can be included in the KS-based upscaling methodology, at the expense of having to solve a larger system of

Table 2. Step-by-step KS approach to the upscaling problem in practice

i. Evaluate the relevant $\mathscr{G}$-KB and $\mathscr{S}$-KB available, and derive the associated vector $\mathbf{g}$ and operator $\Theta_{\mathscr{K}}$.
ii. Formulate the corresponding system of Eqs. (9), and solve it with respect to coefficients vector $\boldsymbol{\mu}$.
iii. Use the $\boldsymbol{\mu}$-values obtained above to compute the operators $\Lambda_{\mathscr{K}}[\zeta_{\alpha,j}]$ and $\Lambda_{\mathscr{K}}[\kappa_\alpha, \zeta_{\alpha,j}]$.
iv. Substitute the last two operators in Eq. (11) to compute the EHC-values, $K_{\mathrm{eff},\alpha,i}$.

equations. It is worth-noticing that KS upscaling derives previous stochastic upscaling results as its special cases under limiting conditions on the information and modelling assumptions used – a fact that demonstrates the generalization power of the KS approach. Below we examine two numerical examples: Example 1 assumes a homogeneous spatial variation of conductivity and offers a simple numerical demonstration of the theoretical fact that the KS approach applies in situations where many traditional upscaling techniques do not. Example 2 considers a non-homogeneous conductivity variation across space.

Example 1: Consider the simple case of unidirectional flow in a spatial domain which has been studied, e.g., by Bakr et al. (1978). The spatial conductivity field $K(s)$ follows a log-normal law so that $\ln K(s) = \Phi(s) = \overline{\Phi(s)} + \varphi(s)$, where $\overline{\Phi(s)} = \bar{\Phi}$ is the known mean of the random log-conductivity field and $\varphi(s)$ is its spatially homogeneous fluctuation with known variance σ_φ^2. The flow BC $J(0) = J_0$ is random (with known mean $\overline{J_0}$ and variance $\sigma_J^2 = \overline{J_0}^2[\exp(\sigma_\varphi^2) - 1]$) and the flow flux q is deterministic. Under these rather limiting conditions, the traditional stochastic approach yields the following EHC,

$$K_{\text{eff}}^{\text{ta}} = \exp[\bar{\Phi} - \frac{1}{2}\sigma_\varphi^2] \tag{12}$$

(the superscript "ta" means "traditional approach"). For numerical illustration, let $\overline{K} = 2.7871$ and $\sigma_K^2 = 0.3983$, in which case $K_{\text{eff}}^{\text{ta}} = 2.6512$. In view of the flow BC and conductivity statistics available the KS approach yields

$$\boldsymbol{\mu}^{\text{T}}\mathbf{g} = \mu_0 + \mu_1\kappa + \mu_2\kappa^2 + \mu_3\zeta + \mu_4\zeta^2 + \mu_5\kappa\zeta \ , \tag{13}$$

where the μ_i-coefficients $(i = 0, 1, \ldots, 5)$ are the solutions of the corresponding system of Eqs. (9), which in this case reduces to the system

$$\left.\begin{array}{llll} \Delta_{\mathscr{K}}[\kappa\zeta] = 0, & \Delta_{\mathscr{K}}[\kappa] = 0, & \Delta_{\mathscr{K}}[\kappa^2] = 0 & \\ \Delta_{\mathscr{K}}[\zeta] = 0, & \Delta_{\mathscr{K}}[\zeta^2] = 0, & \Delta_{\mathscr{K}}[1] = 0, & \Lambda_{\mathscr{K}}[1] = 1 \end{array}\right\} , \tag{14}$$

and $\Theta_{\mathscr{K}} = \Theta_{\mathscr{S}}$. Note that some reformulation of the KS equations can be made when the physical circumstance makes it necessary to replace the conductivity field in Eq. (13) with the corresponding log-conductivity field. On the basis of Equations. (13) and (14), a number of numerical experiments were performed and interesting conclusions were drawn, as follows:

a. In the first numerical experiment we assumed that a case-specific datum in the form of the pdf $f_{\mathscr{S}}(\boldsymbol{\kappa})$ is available for the conductivity field (say, a uniform law with mean 1 and variance 0.333). As we saw in Sect. 3.2, the KS approach is based on a conceptual structure that can go well beyond the limits of the traditional upscaling techniques and can blend different states of core knowledge and uncertain information sources. In this case, KS yielded $K_{\text{eff}}^{\text{ks}} = 2.7723$ (the superscript "ks" denotes KS), which differs from the $K_{\text{eff}}^{\text{ta}}$ by 4.6%. This is expected, since KS processed case-specific information that the traditional technique did not. The difference in EHC values is rather small, since the case-specific information does not differ much from core knowledge about conductivity (larger differences are expected when case-specific information differs considerably from core knowledge). Fig. 1 shows that an error in our knowledge of flow BC has no effect on the $K_{\text{eff}}^{\text{ks}}$ estimate, since the new

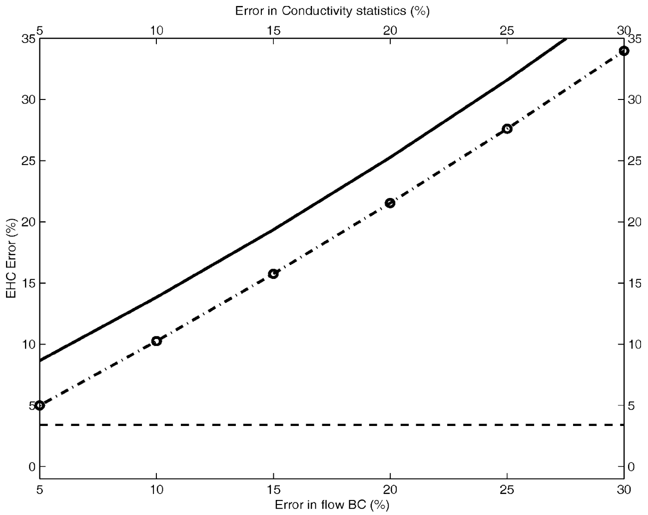

Fig. 1. Plots of error in the KS-based EHC estimate, $K_{\mathrm{eff}}^{\mathrm{ks}}$, due to errors in flow BC (low part dashed line) and in conductivity statistics (continuous line). The error in the traditional EHC estimate, $K_{\mathrm{eff}}^{\mathrm{ta}}$, vs. the error in conductivity statistics (circles) is also plotted for comparison. A $\mathscr{S}$-KB is assumed available in the form of $f_{\mathscr{S}}(\kappa)$

information did not concern the flow BC. On the other hand, a plot of possible errors in our knowledge of the conductivity statistics (%) vs. the resulting errors in the $K_{\mathrm{eff}}^{\mathrm{ks}}$ estimates (also shown in Fig. 1) reveals that the latter increase proportionally to the former. In the same figure a plot of the errors in our knowledge of the conductivity statistics (%) vs. the corresponding $K_{\mathrm{eff}}^{\mathrm{ta}}$ errors is shown for comparison. The error variation in the calculation of the $K_{\mathrm{eff}}^{\mathrm{ta}}$-values is consistently different from the error variation in the calculation of the $K_{\mathrm{eff}}^{\mathrm{ks}}$-values.

b. To compare the $K_{\mathrm{eff}}^{\mathrm{ta}}$-value above with the EHC obtained from the KS upscaling approach under the same flow conditions, we assumed that no $\mathscr{S}$-KB is available and, thus, it is not necessary to revise the initial probabilistic description of flow. Accordingly, the initial values for the μ_i-coefficients are selected so that they are consistent with the same flow BC and conductivity statistics as the traditional method. In this case, it is found that $K_{\mathrm{eff}}^{\mathrm{ks}} = 2.6511$, i.e. the same result as the EHC value obtained by the traditional stochastic technique, as should be expected. Hence, the KS approach exactly reproduces the EHC value of the traditional technique under the same conditions. As Fig. 2 depicts, an error in the flow BC has no effect on the resulting $K_{\mathrm{eff}}^{\mathrm{ks}}$-value. In other words, even when our knowledge of the flow BC is incomplete, the $K_{\mathrm{eff}}^{\mathrm{ks}}$ still produces the theoretical EHC result. This situation is in agreement with the fact that the analytical expression of the $K_{\mathrm{eff}}^{\mathrm{ta}}$ is, in this case, independent of the flow BC [see, Eq. (12)]. Moreover, if our knowledge of the conductivity statistics includes some error, the resulting $K_{\mathrm{eff}}^{\mathrm{ks}}$ value should also include an error. For numerical illustration, a plot of the possible errors in our knowledge of the conductivity statistics (%) vs. the resulting errors in the $K_{\mathrm{eff}}^{\mathrm{ks}}$ estimates is also depicted in Fig. 2 (the plot shows a rather linear variation). In the same figure the corresponding errors in the $K_{\mathrm{eff}}^{\mathrm{ta}}$ calculation are plotted for comparison. Obviously, the error $K_{\mathrm{eff}}^{\mathrm{ta}}$ variation is in close agreement with the error variation in the $K_{\mathrm{eff}}^{\mathrm{ks}}$ calculation above.

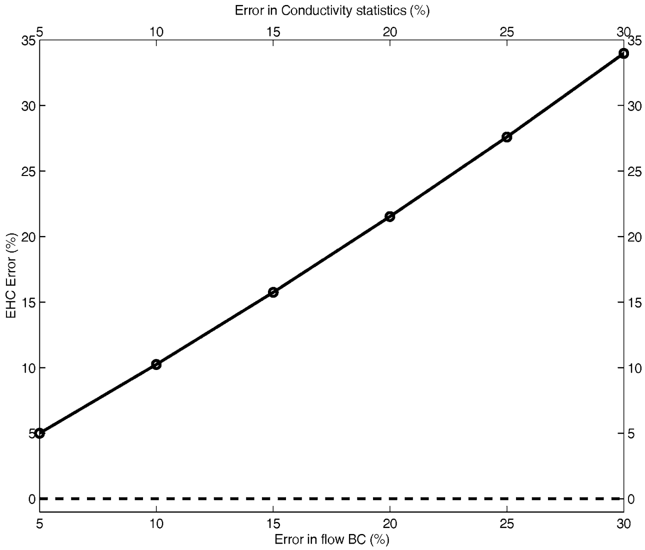

Fig. 2. Plots of error in the KS-based EHC estimate, $K_{\text{eff}}^{\text{ks}}$, due to errors in flow BC (low part dashed line) and in conductivity statistics (continuous line). The error in the traditional EHC estimate, $K_{\text{eff}}^{\text{ta}}$, vs. the error in conductivity statistics (dashed-dotted line with circles) is also plotted for comparison. No $\mathscr{S}$-KB is assumed available

Example 2: The flow and conductivity BC, $J(0) = J_0$ and $K(0) = K_0$, as well as the flow flux q are all deterministic. The random log-conductivity fluctuation field $\varphi(s)$ has a space-dependent variance of the form $\sigma_\varphi^2(s) = c[1 + \text{erf}(s)]$, where c is a constant and "erf" denotes the error-function. Note that fixing $\varphi(s)$ at $s = 0$ has made it a conditioning point and, as a result, the spatial homogeneity hypothesis of Example 1 has been violated. Under these conditions, the traditional stochastic EHC expression is a function of the spatial coordinate s (Oliver and Christakos, 1996),

$$K_{\text{eff}}^{\text{ta}}(s) = \exp\left[\overline{\Phi(s)} - \frac{1}{2}\sigma_\varphi^2(s)\right] \quad . \tag{15}$$

For numerical illustration we select the values $\overline{\Phi(s)} = \bar{\Phi} = 1$ and $c = 0.05$. Then, the coefficients in Eq. (13) are functions of space, i.e. $\mu_i(s)$, and the estimated $K_{\text{eff}}^{\text{ta}}(s)$ is plotted in Fig. 3 vs. the spatial coordinate s. The KS-based EHC estimates across space, $K_{\text{eff}}^{\text{ks}}(s)$, are plotted in the same figure for comparison. Clearly, the $K_{\text{eff}}^{\text{ta}}(s)$ and $K_{\text{eff}}^{\text{ks}}(s)$ plots are in very good agreement with each other (the maximum difference between the two plots is less than 1% and is rather due to numerical error).

The examples above offer valuable insight into the workings of the KS upscaling approach, although considerable work remains to be done in order to test the applicability of KS upscaling in more complex real-world applications. KS may be viewed as a general and flexible upscaling approach that produces complete probability distributions of porous media flow processes using conceptual blending principles, the choice of which depends on the logical and physical features of the situation. These principles establish a dialectic between site-specific data and theory leading to useful EHC estimates.

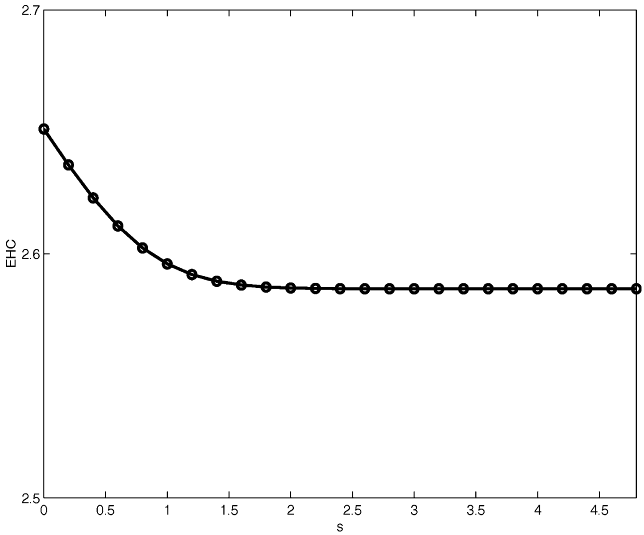

Fig. 3. Plots of traditional EHC $K_{\text{eff}}^{\text{ta}}(s)$ (circles) and KS-based EHC $K_{\text{eff}}^{\text{ks}}(s)$ (continuous line) as functions of space

4 Conclusions

One can find in the literature several upscaling techniques that are based on empirical investigations and formal analysis. A brief review of the various groups of techniques developed over the last few decades was given. The EHC solutions obtained by these groups are based on form approaches and traditional reasoning. The form approaches often do not improve our understanding of the conceptual blending processes at work, whereas conceptual difficulties are, in general, more serious than empirical anomalies. In many cases the problem statement itself dictates the upscaling solution, which makes it hard to expand the possible solutions. The present work emphasizes the decisive role which epistemic arguments enjoy in the rational development of hydrologic sciences, in general, and of the upscaling problem, in particular. The proposed KS-based methodology is an attempt to depart from the traditional way of banausic, formal reasoning underlying many of the existing upscaling techniques. By revising the upscaling fundamentals in the light of the conceptual blending desiderata, one can considerably expand the possible solutions and make finding a solution more interesting. By way of a coda, among the main theoretical and practical advantages of the KS-based upscaling approach over the traditional techniques are the following: (1) It is more open, i.e., it assimilates several types of KB and uncertain information sources relevant to porous media and flow conditions. (2) It is more flexible, e.g., it accounts for conductivity and gradient statistics in a way that does not require an explicit involvement of the statistics in the KS equations. (3) It is more general, i.e., is avoids restrictive model assumptions (e.g., uniformity, lognormality, unbounded domain, uniform flow) and low-order approximations. (4) It is more informative, i.e., it generates complete probability distributions of conductivity-hydraulic gradient across space, instead of single realizations. (5) It is more nested, i.e., it derives many previous results as special cases under limited conditions of porous media flow within the more general KS framework. We conclude by expressing the hope that the present work will direct further attention to the methodological exploration of the upscaling problem at the length and the detail that it deserves. In the end, a researcher only proposes but it is reality that disposes.

References

Bakr AA, Gelhar LW, Gutjahr AL, McMillan JR (1978) Stochastic analysis of spatial variability in sub-surface flows. 1. Comparison of one-dimensional and three-dimensional flow. Water Resour. Res. 14: 263–271

Christakos G, Miller CT, Oliver DL (1993) Stochastic perturbation analysis of groundwater flow. Spatially variable soils, semi-infinite domains and large fluctuations. Stoch. Hydrol. and Hydraul. 7(3): 213–239

Christakos G, Hristopulos DT, Miller CT (1995) Stochastic diagrammatic analysis of groundwater flow in heterogeneous porous media. Water Resour. Res. 31(7): 1687–1703

Christakos G (2000) Modern Spatiotemporal Geostatistics. Oxford Univ. Press, New York, NY

Christakos G (2002a) On the assimilation of uncertain physical knowledge bases: Bayesian and non-Bayesian techniques. Adv. Water Resour 25(8–12): 1257–1274

Christakos G (2002b) On a deductive logic-based spatiotemporal random field theory. Probability Theory and Mathematical Statistics (Teoriya Imovirnostey ta Matematychna Statystyka) 66: 54–65

Christakos G, Bogaert P, Serre ML (2002) Temporal GIS. With Cd-Rom. Springer-Verlag, New York, NY

Cushman JH (1986) On measurement, scale, and scaling. Water Resour. Res. 22(2): 129–134

Cushman JH, Bennethum LS, Hu BX (2002) A primer on upscaling tools for porous media. Adv. Water Resour 25(8–12): 1043–1067

Dagan G (1989) Flow and Transport in Porous Formations, Springer-Verlag, Berlin

Desbarats AJ (1992) Spatial averaging of hydraulic conductivity in three-dimensional heterogeneous porous media. Math. Geology 24(3): 249–267

Deutsch CV (1989) Calculating effective absolute permeability in sandstone/shale sequences. SPE Formation Evaluation, 4(3): 343–348

Durlofsky LJ (1991) Numerical calculation of equivalent grid block permeability tensors for heterogeneous porous media. Water Resour. Res. 27(5): 699–708

Gelhar LW, Axness CL (1983) Three-dimensional stochastic analysis of macrodispersion in aquifers. Water Resour. Res. 19(1): 161–180

Hristopulos DT, Christakos G (1997a) A variational calculation of the effective fluid conductivity of heterogeneous media. Physical Review E 55(6): 7288–7298

Hristopulos DT, Christakos G (1997b) Diagrammatic theory of nonlocal effective hydraulic conductivity. Jour. Stoch. Hydr. Hydraul. 11(5): 369–395

Hristopulos DT, Christakos G (1999) Renormalization group analysis of permeability upscaling. Stoch. Enviro. Res. Risk Assess. 13(2): 131–160

Kitanidis PK (1990) Effective hydraulic conductivity for gradually varying flow. Water Resour. Res. 26(6): 1197–1208

Kolovos A, Christakos G, Serre ML, Miller CT (2002) Computational BME solution of a stochastic advection-reaction equation in the light of site-specific information. Water Resour. Res. 38(12): 1318–1334

Neuman SP, Orr S, Levin O, Paleologos E (1992) Theory and high-resolution finite element analysis of 2-D and 3-D effective permeabilities in strongly heterogeneous porous media. In: Russell TF, Ewing RE, Brebbia CA, Gray WC, Pinder GF (eds.) Mathematical Modeling in Water Resources, Elsevier Appl. Sci. 2: 118–136 New York, NY

Neuman SP, Orr S (1993) Prediction of steady-state flow in nonuniform geologic media by conditional moments: Exact nonlocal formalism, effective conductivities, and weak approximation. Water Resour. Res. 29(2): 341–364

Oliver LD, Christakos G (1996) Boundary condition sensitivity analysis of the stochastic flow equation. Adv. Water Resour. 19(2): 109–120

Rubin Y, Gomez-Hernandez JJ (1990) A stochastic approach to the problem of upscaling of conductivity in disordered media: Theory and unconditional numerical simulations. Water Resour. Res. 26(6): 691–701

Scheibe T, Yabusaki S (1998) Scaling of flow and transport behavior in heterogeneous groundwater systems. Adv. Water Resour. 22(3): 223–238
Serre ML, Christakos G (1999) Modern Geostatistics: Computational BME in the light of uncertain physical knowledge-The Equus Beds study. Stoch. Envir. Res. Risk Assess. 13(1): 1–26

Appendix A

A main function of stochastic porous media upscaling is the calculation of the flow statistics, i.e. the conductivity-gradient cross-covariance $\overline{K_\alpha J_{\alpha,i}}$, in a way that accounts for the $\mathcal{G}$- and $\mathcal{S}$-KB (Sect. 2). In the KS context, the assimilation of $\mathcal{G}$-KB leads to the $\mathcal{G}$-equations (also called, teleologic equations)

$$\left.\begin{array}{l} K_{\mathrm{eff},\alpha,i}\overline{J_{\alpha,i}}\Big|_{\mathcal{G}} = \int \mathrm{d}\boldsymbol{\kappa} \int \mathrm{d}\boldsymbol{\zeta} \kappa_\alpha \zeta_{\alpha,i} f_{\mathcal{G}}(\boldsymbol{\kappa},\boldsymbol{\zeta}), \quad \overline{J^\lambda_{\alpha,i}}\Big|_{\mathcal{G}} = \int \mathrm{d}\boldsymbol{\kappa} \int \mathrm{d}\boldsymbol{\zeta} \zeta^\lambda_{\alpha,i} f_{\mathcal{G}}(\boldsymbol{\kappa},\boldsymbol{\zeta}) \\ \beta^{(\boldsymbol{\rho})}_{\boldsymbol{\alpha}} = \int \mathrm{d}\boldsymbol{\kappa} \int \mathrm{d}\boldsymbol{\zeta} \kappa^{\rho_1}_{\alpha_1} \kappa^{\rho_2}_{\alpha_2} \cdots \kappa^{\rho_q}_{\alpha_q} f_{\mathcal{G}}(\boldsymbol{\kappa},\boldsymbol{\zeta}), \quad 1 = \int \mathrm{d}\boldsymbol{\kappa} \int \mathrm{d}\boldsymbol{\zeta} f_{\mathcal{G}}(\boldsymbol{\kappa},\boldsymbol{\zeta}) \end{array}\right\} \tag{A.1}$$

(for all $i, \lambda, \boldsymbol{\alpha}, \boldsymbol{\rho}$), where $f_{\mathcal{G}}(\boldsymbol{\kappa},\boldsymbol{\zeta})$ has the form of Eqs. (5) so that the vectors $\boldsymbol{\mu}$ and $\mathbf{g}$ are consistent with mean flow law (1), associated BC, and conductivity statistics. In light of $\mathcal{S}$-KB, an updated system of upscaling equations is derived in the integration stage:

$$\left.\begin{array}{l} K_{\mathrm{eff},\alpha,i}\overline{J_{\alpha,i}}\Big|_{\mathcal{K}} = \int_{I_\kappa} \mathrm{d}\boldsymbol{\kappa} \int_{I_\zeta} \mathrm{d}\boldsymbol{\zeta} \kappa_\alpha \zeta_{\alpha,i} f_{\mathcal{K}}(\boldsymbol{\kappa},\boldsymbol{\zeta}), \quad \overline{J^\lambda_{\alpha,i}}\Big|_{\mathcal{K}} = \int_{I_\kappa} \mathrm{d}\boldsymbol{\kappa} \int_{I_\zeta} \mathrm{d}\boldsymbol{\zeta} \zeta^\lambda_{\alpha,i} f_{\mathcal{K}}(\boldsymbol{\kappa},\boldsymbol{\zeta}) \\ \beta^{(\boldsymbol{\rho})}_{\boldsymbol{\alpha}}\Big|_{\mathcal{K}} = \int_{I_{\boldsymbol{\kappa}}} \mathrm{d}\boldsymbol{\kappa} \int_{I_\zeta} \mathrm{d}\boldsymbol{\zeta} \kappa^{\rho_1}_{\boldsymbol{\alpha}_1} \kappa^{\rho_2}_{\boldsymbol{\alpha}_2} \cdots \kappa^{\rho_q}_{\boldsymbol{\alpha}_q} f_{\mathcal{K}}(\boldsymbol{\kappa},\boldsymbol{\zeta}), \quad 1 = \int_{I_{\boldsymbol{\kappa}}} \mathrm{d}\boldsymbol{\kappa} \int_{I_\zeta} \mathrm{d}\boldsymbol{\zeta} f_{\mathcal{K}}(\boldsymbol{\kappa},\boldsymbol{\zeta}) \end{array}\right\} \tag{A.2}$$

(for all $i, \lambda, \boldsymbol{\alpha}, \boldsymbol{\rho}$), where $f_{\mathcal{K}} = f^{\mathrm{bc}}_{\mathcal{K}}$ or $f^{\mathrm{mb}}_{\mathcal{K}}$ is the pdf of Eqs. (6) or (7). By combining Eqs. (A.1) and (A.2) and in view of Eqs. (8) and (10) we obtain Eqs. (9) which can be solved for $\mu_{\boldsymbol{\alpha}}$ [as already mentioned, the BC assumed for law (1) can affect the initial $\mu_{\boldsymbol{\alpha}}$-values]. Subsequently, the $K_{\mathrm{eff},\alpha,i}$ values are obtained from Eq. (11).