Materials and Structures (2015) 48:771–796
DOI 10.1617/s11527-014-0515-0

ORIGINAL ARTICLE

Optimization method, choice of form and uncertainty quantification of Model B4 using laboratory and multi-decade bridge databases

Roman Wendner · Mija H. Hubler · Zdeněk P. Bažant

Received: 5 June 2014 / Accepted: 19 December 2014 / Published online: 8 January 2015

Abstract The preceding article describes a new multi-decade creep and shrinkage prediction model, labeled B4, which extends and improves the previous RILEM recommendation B3, and a separate article presents a new large worldwide database of laboratory data on creep, drying shrinkage and autogenous shrinkage. This article presents the general optimization concepts used to verify and calibrate Model B4. The main objective is multi-decade, even 100-year, prediction, which is what is needed for sustainable design of long-span bridges, tall buildings and other large concrete structures. Since the existing worldwide database is insufficient by far for purely experimental verification and calibration of multi-decade creep, the importance of choosing a model form that is theoretically and physically justified is emphasized. So is the choice of a model form that is able to fit closely individual broad-range creep and shrinkage curves for one and the same concrete, which are free of the huge obfuscating scatter due to differences in concrete composition. The development and calibration of the formulae for predicting the parameters of the creep and shrinkage equations from the composition and strength of concrete is described. An effective method for statistical optimization of the fits of a new database comprising thousands of laboratory test curves is presented. Various types of bias in the database are counteracted by data weighting. To predict and calibrate multi-decade creep, a method to combine the laboratory database with a database of excessive multi-decade deflections of large span bridges is outlined. This leads to a significant increase of the slope of the terminal trend of predicted creep in a logarithmic plot. Finally, the statistical approaches for using the hybrid laboratory-bridge database for multi-objective fit optimization and for Bayesian updating are explained and discussed.

Keywords Optimization · Database · Model development · Creep · Shrinkage · Statistical bias

R. Wendner · M. H. Hubler
Civil and Evironmental Engineering, Northwestern University, Evanston, IL, USA

Present Address:
R. Wendner
Christian–Doppler Laboratory on Life−Cycle Robustness of Fastening Technology, Institute of Structural Engineering, University of Natural Resources and Life Sciences, Vienna, Austria

Present Address:
M. H. Hubler
Civil and Environmental Engineering, Massachusettes Institute of Technology, Cambridge, MA, USA

Z. P. Bažant (✉)
Civil and Mechanical Engineering and Materials Science, Northwestern University, Evanston, IL, USA
e-mail: z-bazant@northwestern.edu

1 Introduction

The design of new, and assessment of, existing concrete structures requires the prediction of structural

response under mechanical and environmental loads for the full service life. Bridges and other important structures are typically designed for a lifetime of at least 50 years, but large structures usually 100 years or more. Such design obviously requires realistic multi-decade prediction models. The prediction models that are currently endorsed by engineering societies suffer from a series of shortcomings which reduce the safety margin in various situations and compromise the economic maintenance of the built infrastructure. These are primarily the purely empirical formulations, calibrated by limited data of insufficient duration. Often they give incorrect predictions of the true behavior of concretes, especially the modern, high performance, ones. Rare though the disasters attributable to poor creep and shrinkage prediction are, the costs of excessive deflections of large bridges and of their retrofits are enormous.

When the experimental data uniformly cover the entire range of practical interest, a purely empirical approach, with no theory, can often be used to obtain a satisfactory model. This has been the case for many specifications of the concrete design codes. However, there are two instances where sufficient data do not exist and can hardly be obtained—the size effect on structural strength, discussed elsewhere, and the multi-decade creep, which is in the focus here.

In the largest existing database, comprising thousands of creep and shrinkage tests [1], 95 % of the tests have durations not exceeding 6 years, and a 12-year duration is exceeded only by a few tests (some of which suffered from lapse of long-term control). In such a case (as well as in the case of size effect), the model adopted must have a sound theoretical basis.

For multi-decade extrapolations, the model may be calibrated by inverse analysis of relative multi-decade excessive creep deflections of bridges. The bridge data can be taken into account in two ways—either by simultaneous optimization of the fits of combined laboratory and bridge databases, or by Bayesian updating of the model.

The objective of this article is to describe the methods used and provide the background behind the development.

The next two articles [2, 3] will employ these methods to calibrate and verify the creep and shrinkage parts of Model B4 and compare them statistically to other models.

2 Requirements for model form to be optimized

2.1 Theoretical requirements

Optimizing a model of a purely empirical or intuitive form cannot be expected to provide good extrapolations to multi-decade or even century long durations. Same as model B3, the 1995 RILEM recommendation for creep and shrinkage predictions [4], here we take the viewpoint that the model to be optimized should have a form based on an existing undisputed theory with physical foundations. Even though such a physically based formulation does not automatically ensure obtaining a set of correct parameters from a close data fit, the probability of arriving at a poor model providing systematically erroneous extrapolations is significantly lower than it is for an empirical model. In an empirical model, a distortion of model parameters due to the compensation of errors is likely [5]. At the present level of understanding, the form of a physically based creep model should have the following attributes:

1. It should be based on the solidification theory [6]. This theory is the only existing thermodynamic basis for aging viscoelasticity of a solidifying material, in which the solid skeleton is stiffening because of gradual deposition of unstressed solid on the pore surfaces, precipitating from a solution. The point to note is that thermodynamics cannot be formulated for a material with time-variable properties. The only way to formulate thermodynamics of an aging and hence stiffening material is to consider the changing concentration of a material constituent that itself is non-aging.
2. The solidification theory directly yields the rate of compliance rather than the compliance itself [6]. The fact that the compliance rate obtained cannot be integrated to a simple closed-form expression for the compliance is a mere loss of beauty but represents no impediment to practical applications since the rate form suffices for computations. The microprestress-solidification theory [7, 8] further indicates that the long-term asymptotic trend of creep should be logarithmic (indeed, the available test data give no hint of an approach to some final asymptotic value and support the logarithmic form of the terminal trend [9]).
3. The compliance form should satisfy the condition of non-divergence of compliance curves for different ages at loading. Many models, and all

those considered later in this article, except B3 and B4, violate this condition for some periods of time. This leads to two objectionable features:

(a) The creep recovery curves obtained from these models according to the principle of superposition can be non-monotonic, i.e., give a recovery reversal for some combinations of loading and unloading times [10], see also [3].
(b) For long enough times, the relaxation curves calculated from these models by the principle of superposition cross into the opposite stress sign [10].

Based on the solidification theory and on the fact that aging viscoelasticity must be describable by the Maxwell chain model, the aforementioned two features are thermodynamically impossible. Besides, none of them have been verified experimentally.

4. Because the drying shrinkage is roughly proportional to water loss caused by diffusion through the pores, the average drying shrinkage in a cross section of a concrete structural member must conform to the basic general characteristic properties of the diffusion theory for moisture in concrete [11]. Consequently, the shrinkage strain should begin to evolve as a square root of time after the drying exposure and should terminate by an exponential of time square root to a constant final asymptotic value. The shrinkage half-time must scale with the square of effective thickness or volume–surface ratio of the cross section, and the factor of this ratio should vary with the cross-section shape according to the diffusion theory.
5. The drying creep has the same physical source as shrinkage. Thus, it should exhibit the similar properties as mentioned above for shrinkage and should be formulated as a separate term of the compliance function.
6. The physical source of autogenous shrinkage is different from drying shrinkage and, consequently, it should be modeled by a separate term with no diffusion characteristics.

The previous model B3 was constructed to satisfy all these requirements except the last. Model B4 now satisfies them all, including the last. Compared to other models, this improves the chances that multi-decade creep and shrinkage extrapolations would be realistic.

Looking at the foregoing listing of various properties, one must not get the impression that all the physical mechanisms are known. To the contrary, the mechanisms of creep and shrinkage on the nanoscale of the calcium–silicate–hydrate (C–S–H) are not fully understood. The mechanism of autogenous shrinkage and its interaction with simultaneous drying has not been clarified. The knowledge of concrete composition effects on the creep and shrinkage is still for the most part empirical. Consequently, a calibration of the model equations by an extensive database is essential.

2.2 Fitting individual test curves

Another requirement for the choice of the mathematical form of the equations to be optimized is that they must be able to fit closely the individual test curves. This requirement is essential. The equation form can experimentally be validated *only* by fitting the time curves of creep or shrinkage from individual tests of long enough duration on one and the same concrete (preferably all the specimens being from the same batch). This must be the first check of the validity of any theoretically based formulation and is imperative for revealing the shortcomings of empirical models (in Sect. 12, this and further aspects will be discussed at length).

Due to the large intrinsic scatter of creep and shrinkage measurements that can be attributed to variations in composition, global statistics such as the coefficient of variation (C.o.V.) of the errors in fitting the database as a whole are rather insensitive to changes in the function form. Figure 1 shows the intrinsic scatter of (a) drying creep, (b) basic creep, and (c) total shrinkage for a subset of the full NU database that is limited to a narrow range of environmental conditions, effective member thickness, D, age of loading, t', and start of drying, t_0. In direct comparison, the significant differences in function shape between the typical predictions according to the new Model B4 [12], the *fib* Model Code 2010 model [13], and the ACI92 model [14], using the mean values of the parameter ranges, loose in importance. If one would fit the compliance curves with superimposed sinusoidal oscillations (obviously nonsensical), the C.o.V. of the errors in fitting the database as a whole would barely change, even if the oscillation amplitude were several times larger than typical model errors of individual curves.

Although only a fraction of the available laboratory creep and shrinkage curves are of sufficient duration,

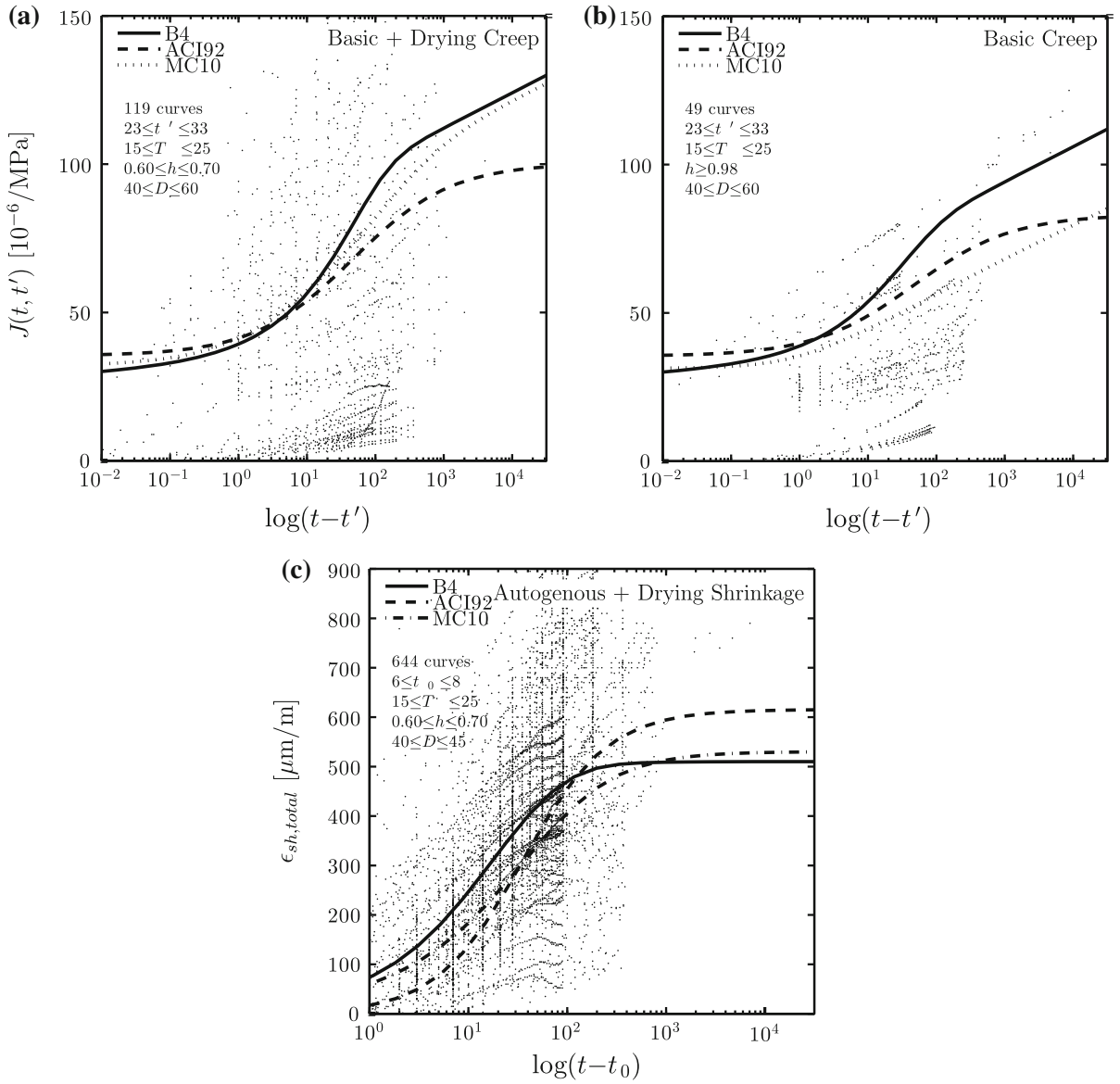

Fig. 1 Intrinsic scatter due to differences in concrete composition versus shape of prediction models; **a** drying creep, **b** basic creep, and **c** total shrinkage

they suffice for checking the equation form [1] if analyzed individually.

2.3 Uncertainty quantification

A good prediction model must provide an estimate of prediction uncertainty. This is needed for the life-time prediction of new structures as well as the reliability assessment of existing structures [15–17]. Apart from presenting the methodology for calibrating the multi-decade creep and shrinkage prediction, an uncertainty quantification will be presented later. The required stochastic models for composition parameters, as well as the default parameter sets determined from strength, are derived in separate articles [2, 3].

The main challenge at hand lies in the extraction of empirical predictor equations for the model parameters based on a limited set of biased and incomplete data for various concrete compositions, environmental conditions and geometry. It is paramount that a good

prediction model for creep and shrinkage be able to accurately capture the most fundamental influence quantities such as temperature T, environmental humidity h, size D, and concrete composition. Yet, only a limited number of parameter combinations have actually ever been experimentally sampled. Thus, the trends for, e.g., size that have been investigated for certain environmental conditions only, have to be generalized to all conditions. Naturally, this introduces large uncertainties. Nevertheless, it represents the only feasible approach to formulate a prediction model with a wide range of applicability and acceptable prediction quality, at least on average.

In addition to a general fitting strategy that has been optimized to deal with the particular challenges of creep and shrinkage prediction, we will introduce a weighting scheme to counteract the inherent bias, formulate suitable statistical indicators for the quality of fit, choose the relevant input quantities, estimate the measurement errors and deal with incomplete information.

It must be emphasized that structural observations, although far from ideal and generally necessitating complex inverse analysis, represent the only viable source of multi-decade information. Relevant are mainly multi-decade data on excessive deflections of long-span prestressed bridges (if the deflection is not excessive, the prestress deflection offsets the gravity deflection and thus nothing can be learned about the creep law). To extract information on creep, the deflection due to multi-decade steel relaxation in tendons must be separated, and sometimes the same needs to be done for cracking. In this regard, long-term monitoring of pure compression members such as arches or columns in tall buildings would seem preferable, especially if the columns are not exposed to climate. The data on tall buildings, however, are even more difficult to obtain than bridge data.

For the actual optimization, two approaches have been considered. One represents a weighted, multi-objective optimization utilizing the full range of the hybrid bridge deflection and laboratory test database, the other a two-phase approach involving a traditional optimization of short-term laboratory data followed by Bayesian updating with structural observations. The latter can be applied not only during model development but also during construction, to update the prediction based either on laboratory tests or on monitored deflections.

3 Model complexity as penalty for characterizing average cross section strain and other obstacles to calibration

3.1 Non-homogeneous state of cross section

Creep and shrinkage should ideally be characterized by a point-wise constitutive equation. However, except for basic creep for which the cross section behavior also represents the constitutive equation, the cross section is in an evolving nonuniform state of stress and pore humidity. Thus, the point-wise constitutive equation would have to be extracted from laboratory data by inverse three-dimensional finite element analysis of test specimens, in which the evolutions of residual stresses, cracking and moisture diffusion in test specimens would have to be calculated.

The point-wise constitutive equation, if identified, would be much simpler since it would depend on a much smaller number of variables. The moisture conditions and the cross-sections size and shape would become the boundary conditions to partial differential equations solved by finite elements. However, whereas such point-wise characterization is doubtless the future, it would require three-dimensional analysis of displacements, stresses, cracking and pore humidities of all creep sensitive structures. The current state of art in structural design is not yet ready for that.

What is needed today for most design practice is a model for the average creep and shrinkage of cross-sections of beams, plates and shells. Such an average model, which is the goal of the present work, has two advantages: (1) it simplifies the task for the designer, as it makes possible a simplified one- or two dimensional analysis of most structures; and (2) it makes it easier to identify the creep and shrinkage model by applying powerful statistical optimization tools to a large database. The optimization would become a preposterous task if inverse finite element analysis were to be conducted for each test specimen.

3.2 Are data on old concretes still relevant?

It is often pointed out that calibration by data for old concretes of lower strength, high water–cement ratios and no admixtures has no relevance for modern concretes. However, this opinion is an exaggeration. The data for old concretes (1980s and earlier) are still very useful, for three reasons:

1. Long-term laboratory tests are available only for these old concretes. The present model could not have been properly calibrated for multi-decade predictions using the data on modern concretes only.
2. Although the data for these old concretes are not relevant for calibrating the effect of concrete composition and strength on creep and shrinkage parameters of modern concretes, they are nevertheless relevant for revealing the long-term shape of creep and shrinkage curves. The reason is that, in old as well as modern concretes, the source of creep is the same—the tricalcium silicate hydrate (C–S–H). Admixtures and additives affect the hydration reactions in a quantitative, rather than qualitative, way, altering the rates and magnitude of creep and causing a 'stiffer' response as though the hydration were accelerated.
3. For drying shrinkage and drying creep of old and new concretes, the source is also the same—the capillary tension, surface tension and spreading pressure in adsorbed water layers, and disjoining pressure in nano-pores of C–S–H which causes the microprestress. Admixtures and additives, to a large part, merely accelerate or delay the formation of the C–S–H micro–structure and thus creep and shrinkage, but the shape of the curves of basic creep, drying creep and drying shrinkage of old and modern concretes appears to be about the same. What is different is the effect of composition and high strength on the creep and shrinkage parameters. Similar comments can be made about various types of cement replacement products, e.g., silica fume.

Note that the foregoing arguments, of course, do not apply to autogenous shrinkage. Its source is the self-desiccation and the volume changes during chemical reactions of additives and admixtures, which are mostly absent from old concretes.

3.3 Database

For the purpose of developing an improved model for (multi-decade) creep and shrinkage prediction, the largest database of laboratory tests in existence has been created [1]. This database, which more than doubles the size of the previous RILEM database [18] (and can be freely downloaded from http://www.civil.northwestern.edu/people/bazant/ and http://www.baunat.boku.ac.at/creep.html) includes approximately 1,400 test curves of creep and almost 1,800 test curves of shrinkage. Among these, almost 800 creep curves and more than 1,000 shrinkage curves (including some that were already part of the RILEM database) contain significant amounts of admixtures.

The development of multi-decade prediction models is seriously impaired by the fact that only 5 % of the creep data are test curves for durations of more than 6 years, and only 3 % for 12 years or more. The latter include the 23-year tests of Troxell et al. [19], the 18-year tests of Russell and Burg [20, 21], the 12-year tests of Browne [22] and of the Bureau of Reclamation [23, 24]; also, there are 30-year tests of Brooks [25], but for those a sudden change of slope in log-time at about 6 years suggests a lapse of environmental control.

Additional information on multi-decade creep is available only through inverse analysis of structural observations. To this end, a database of bridge deflection histories of 69 excessively or strongly deflecting large-span prestressed bridge spans from around the world was assembled at Northwestern University [1, 26]. For most of them, neither the exact geometry nor the prestress details and concrete properties with composition are available, which destroys the chance of full inverse analysis. Yet, statistical analyses revealed a systematic behavior—a linear trend of long-time deflection with the logarithm of time, $\log t$. Although the linearity of this trend in $\log t$ agrees with what was previously deduced from some laboratory data [9], the slope of this trend seen on the bridge data is significantly higher than it is for any current creep prediction model.

4 Essentials of Model B4

The new B4 prediction model, which is the successor of the well-established 1995 RILEM recommendation, Model B3, extends the range of applicability to modern concretes with admixtures, characterized by relatively large autogenous shrinkage. Model B4 captures the behavior of Portland cement concretes and gives formulae for an approximate prediction of model parameters from the composition (Model B4) or strength (Model B4s) of concrete, and the environmental conditions. Compared to Model B3, further improvements include a recalibration for multi-decade behavior

and, importantly, a split of shrinkage strain into a sum of drying shrinkage, $\epsilon_{sh}(t, t_0)$, and autogenous shrinkage, $\epsilon_{au}(t)$, where t is the current time and t_0 is the time at exposure to drying. In the service stress range (up to $0.40f'_c$), a linear dependence of creep strain on stress may be assumed as an acceptable approximation. This means that, for constant uniaxial stress σ applied at age t', the strain evolution is given by

$$\epsilon(t) = \sigma J(t, t') + \epsilon_{sh}(t, t_0) + \epsilon_{au}(t) \quad (1)$$

The form of the compliance function, $J(t, t')$, as introduced in [4] on the basis of the solidification theory, is here adopted without any change. The compliance function is split into the basic creep part C_0 (no moisture exchange) and drying creep part C_d and applies to compressive as well as tensile loading;

$$J(t, t') = \frac{1}{E_0} + C_0(t, t') + C_d(t, t', t_0), \quad (2)$$

where $1/E_0 \approx q_1/E_{28}$ is the asymptotic (truly instantaneous) compliance estimated from the 28-day Young's modulus; q_1 to q_4 are scaling parameters for basic creep, and q_5 = scaling parameter for drying creep.

The differences between tensile and compressive creep that are reported in literature are likely caused by cracking damage which should properly be accounted for in a model of rate-type form. This is nowadays routine in finite element computations. The evolution of damage must not be included in the material model for creep because it is not a constitutive property as is depends on structure geometry, size and loading, and because it affects the diffusion of moisture in the structure, which also depends on structure geometry and size and influences creep. A full inverse analysis of drying creep tests is a highly demanding task that involves two-way coupling with models for material damage, moisture (and heat) transfer, and the C–S–H solidification (or aging). Nevertheless, model B4 approximately characterizes these effects in the sense of an average cross-section behavior of axially loaded structural members. A more accurate characterization requires a point-wise constitutive model and three-dimensional structural analysis, which is still seen as too complex for the current design practice.

The basic creep compliance is the sum of an ageing viscoelastic term (subsequently referred to as the q_2 term), a non-ageing viscoelastic contribution (q_3 term), and a flow term (q_4 term) (see Eq. 27 in [12]), where $Q(t, t')$ = binomial integral given in Eq. 28 in [12];

$$C_0(t, t') = q_2 Q(t, t') + q_3 \ln\left(1 + (t - t')^{0.1}\right) + q_4 \ln\left(\frac{t}{t'}\right) \quad (3)$$

Integral $Q(t, t')$ cannot be evaluated in a closed form, but a very good closed-form approximation exists [6]. In step-by-step computer analysis one needs only the time rate $\partial Q(t, t')/\partial t$ which can be given in a closed form. For the detailed formulation see [12].

The time functions of drying shrinkage and autogenous shrinkage are S-shaped curves bounded by final shrinkage values $\epsilon_{sh\infty}$ and $\epsilon_{au\infty}$ respectively. Accordingly, the time function of drying creep is bounded, too. The basic creep compliance, however, is unbounded, having a finite terminal asymptotic slope $0.1q_3 + q_4$ in the logarithmic time scale.

The fitting of a shrinkage and creep model to data requires separating the contributions from two and four curves, respectively, see Fig. 2b. Because the shapes of these curves are similar, the optimization thus leads to multiple local minima and uncertain convergence.

5 Optimization strategy

The development of an effective strategy for calibrating Model B4 involved various tasks and test types as visualized in Fig. 2. For each concrete composition, one must in theory have a sufficient number of tests revealing the environmental influences, including environmental relative humidity h and temperature T, the boundary conditions, the applied uniaxial stress σ, and the specimen size, characterized by the effective thickness $D = 2V/S$, where V/S is the volume to exposed surface ratio. This already demanding list of tests increases proportionally with the number of different concrete compositions, cement types, including the regular cement, rapid-set cement and slow hardening cement (labeled as R, RS and SL, in agreement with the definition according to *fib* [13]) and admixture types, as well as the aggregate types of which the database [1] contains twelve.

The long-time deformations in Model B4, which consist of (1) autogenous shrinkage ϵ_{au}, (2) drying shrinkage ϵ_{sh}, (3) basic creep σC_0 and 4) drying creep

σC_d, interact as sketched at the bottom of Fig. 2. They cannot be measured separately, with two exceptions—the autogenous shrinkage, which can be measured on load-free sealed specimens, and the basic creep of normal and low strength concretes with high water–cement ratios, which can be measured on loaded sealed specimens because the autogenous shrinkage can be assumed to be negligible. In high strength concretes, every strain measurement includes a significant contribution of autogenous shrinkage. Note that the above statement only holds for the average cross-section response quantities.

Thus, for concretes with significant autogenous shrinkage, the drying shrinkage can be obtained only by inverse analysis of both sealed specimens and specimens exposed to a controlled environment. For the sake of simplicity, models for the mean cross sectional response, including Model B4, require the assumption of additivity of autogenous and drying shrinkage and of basic and drying creep (note that the latter would not be necessary for a point-wise rate-type constitutive relation which can capture cracking and moisture diffusion as separate phenomena). The drying creep tests are the most difficult to analyze as the recorded data include contributions from all the four deformation components.

The present optimization strategy has specifically been conceived to identify and calibrate the creep and shrinkage model for the mean cross-sectional behavior of beams and plates with or without a centric axial (or in-plane) load. In addition to coping with the interaction of deformation components, the non-uniqueness of the time functions and the measurement errors, a joint optimization strategy for the laboratory and structural databases needs to be also formulated.

While the laboratory data provide insight into the effects of composition, environmental conditions, loading parameters and geometry, the only significant source of multi-decade information are the measurements on structures—in this case, the data on excessive deflections of bridges. The heteroscedastic nature [28] of the available data and their statistical sampling bias complicate matters but can be handled by properly formulated statistical metrics and objective functions in combination with a rational weighting scheme [27].

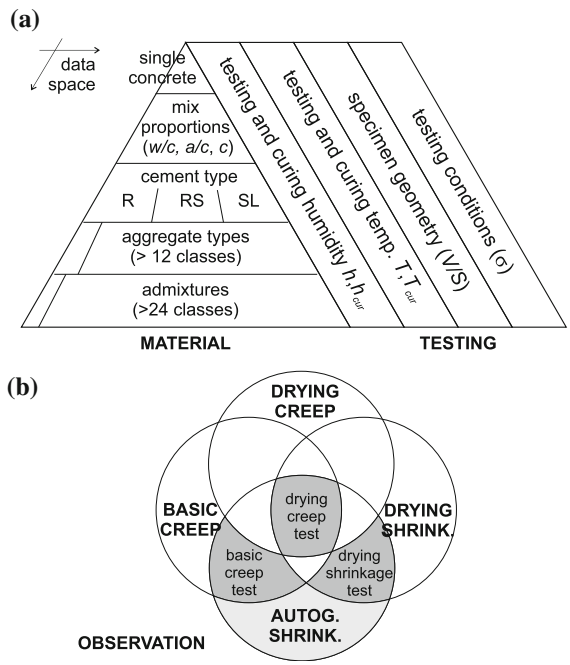

Fig. 2 Complexity of model optimization problem; **a** number of potential tests depending on intrinsic, environmental and load factors; **b** model components versus measurable quantities

6 Objective function and weighting

The calibration of a prediction model can ultimately be reduced to a standard optimization problem if appropriate metrics to handle the heteroscedastic nature of the data [28] and to adequately represent the quality of fit are introduced. The combination of different data sources (here shrinkage strains ϵ, compliance J, and bridge deflections δ) introduces further complications. The data are of different orders of magnitude, exhibit inconsistent sensitivities to environmental and intrinsic material properties, and are associated with varying levels of inherent material uncertainty. On top of that, every data set shows a certain bias toward specific testing conditions or material compositions, which depends on the experimenter's preferences, the need for easy testing, or the preferences of the engineering community. A subjective bias towards certain compositions or environmental conditions is often intended in the sense of an importance weight. A model that is practically relevant may be considered preferable over an objective model that introduces the same degree of error in all cases.

The usual way to deal with heteroscedastic data as well as the inherent bias in the input information is weighting. Although, in a strict statistical sense,

weighting introduces bias, it is the only way to counteract the unwanted pre-existing bias, e.g., bias towards short-term data, small specimen sizes, and low ages at the start of loading or drying. Figures 3 and 4 explain the concept of hyperbox weighting introduced in [29] for the calibration of shrinkage and creep models, respectively. The term hyper-box refers to the *n*-dimensional hyper-space of the optimization problem in which all the parameters, such as time, composition, geometry, environmental conditions, should in theory be covered by data uniformly, in order to attain an unbiased ideal model.

The individual weights w_{ij} of curve i and half-decade j (delimited in the logarithmic scale) counteract the bias stemming from the data point density or a preference towards a certain data range. The weights W_i are intended to remove bias due to material composition or preferred testing conditions. They also allow introducing importance weights, either for experimenters that are known to be particularly thorough or for types of tests that are particularly relevant for the model development and its future application.

Additionally, weights may be assigned due to smoothness and consistency of a particular data set, quality of data reporting, and especially the completeness, e.g., whether or not the full time range from the onset of drying shrinkage to the asymptotic approach to the final value is covered. As discussed in [1], a shrinkage strain or compliance curve that was re-digitized from the article in which the curve was (unfortunately) reported in a linear scale may introduce short-term data errors exceeding 100 %.

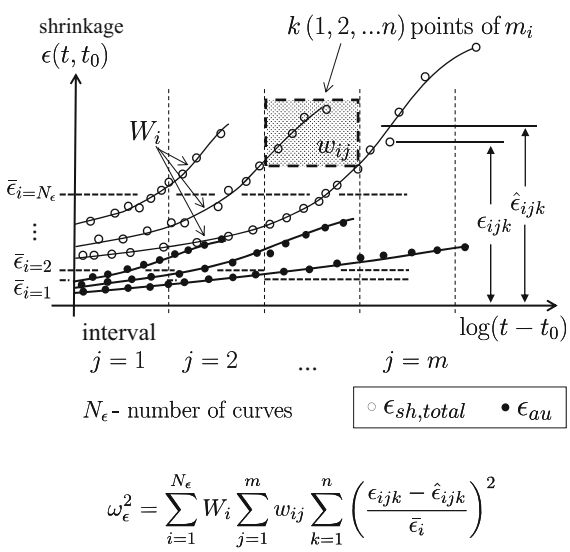

Fig. 3 A sketch of the weights assigned according to hyperboxes to avoid testing bias in the shrinkage data. Weights are assigned to each region as defined by the associated composition, testing, time interval, and importance values such that no set of points dominates the residual measure for optimization

A statistically suitable measure to describe the quality of fit for data of heteroscedastic nature, and thus the basis of the objective function, is the C.o.V., ω. It represents the dimensionless ratio of the root mean square error (RMSE) to the data mean $\bar{y}$, i.e.:

$$\omega = \mathrm{RMSE}/\bar{y} \tag{4}$$

Thus, the total objective function value with contributions of the single data sources is given by

$$\omega = \sqrt{\sum_{i=1}^{N_y} W_i \sum_{j=1}^{m} w_{ij} \sum_{k=1}^{n} \left(\frac{y_{ijk} - \hat{y}_{ijk}}{\bar{y_i}} \right)^2}, \tag{5}$$

where y_{ijk} is the kth measured value in half-decade j on curve i; $\hat{y}_{ijk}$ is the predicted value of point ijk, and $\bar{y_i}$ is the mean of the measured data of curve i. The total sum of weights of all N_y curves with N points must be normalized to 1, i.e.,

$$\sum_{i=1}^{N_y} W_i \sum_{j=1}^{m} \sum_{k=1}^{n} w_{ij} = 1 = \sum_{k=1}^{N} w_k \tag{6}$$

For convenience, the weights W_i and w_{ij} may be combined into independent weights w_k assigned directly to all N data points of any given data source. The optimization problem finally reads:

$$\hat{\mathbf{X}} = \begin{cases} \min(W_\epsilon \omega_\epsilon^2) & \text{shrinkage} \\ \min(W_J \omega_J^2 + W_\delta \omega_\delta^2) & \text{creep} \end{cases} \tag{7}$$

with $W_\epsilon = 1$ is the total weight of the shrinkage laboratory database, W_J is the total weight of the creep laboratory database, and $W_\delta = 1 - W_J$ is the total weight of the bridge deflection database. During the optimization of the creep model, the total contributions of the laboratory database and the bridge database, collecting bridge deflections, were weighted as 2:1. That means with N_J laboratory data sets and N_δ bridge deflection records that $W_J \sum_{k=1}^{N_J} w_k = 2W_\delta \sum_{k=1}^{N_\delta} w_k = 2/3$.

Using the bridge deflection records jointly with the laboratory compliance data is, unfortunately, not

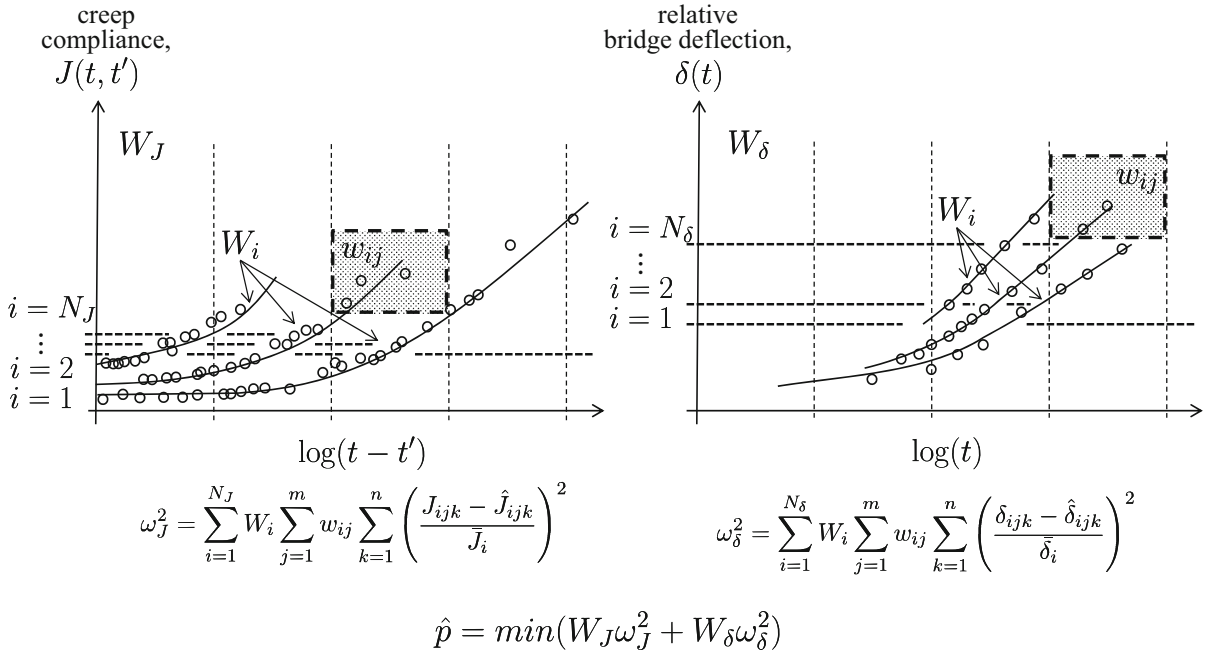

Fig. 4 A sketch of the weights assigned according to hyperboxes to avoid testing bias in the laboratory creep data and relative bridge deflections. Weights are assigned to each region as defined by the associated composition, testing, time interval, and importance values such that no set of points dominates the residual measure for optimization. Additional weights are given to each of the two databases to counteract the abundance of short term measurements for the joint optimization

straightforward. Either bridge deflections $\delta(t)$ must be converted to compliance $J(t,t')$, or the predicted compliance function transformed to an evolution of deflections. Only then, the residuals can be calculated and only then a consistent contribution to the objective function can be formulated. For an accurate conversion, a complex inverse analysis using three-dimensional finite element simulations like those in [30] would be required.

However, if the deflection $\delta(t)$ is known at a certain reference time $t = t_{\text{ref}} \gg t_{\text{bc}}$ where t_{bc} is the average concrete age at span closing, the deflection trend can simply be extrapolated to long times by assuming similarity to the compliance function. As shown in [26], for $t \geq t_{\text{ref}}$ the age and slab thickness differences among the box girder segments may be ignored and the age of the concrete may be characterized by one common effective (or average) age t_{bc} [26].

Further necessary simplifications include the definition of one common effective age t_{a} at which the self-weight bending moments are introduced in the erected cantilever, instead of considering the gradual increase of the bending moment during erection of the bridge. For the purpose of this investigation the values t_{bc} is the 120 days and $t_{\text{a}} = 60$ days were used. This approximation yields acceptable results only for sufficiently large reference times (such as $t_{\text{ref}} \approx 1{,}000$ days since the closing of the span), which ensure the complex initial behavior due to the effects of drying and construction sequence and to the differences in age to have almost died out. The extrapolation formula of the deflection from t_{ref} thus reads:

$$\frac{\Delta J(t,t_{\text{a}})}{\Delta J(t_{\text{ref}},t_{\text{a}})} \approx \frac{J(t,t_{\text{a}}) - J(t_{\text{bc}},t_{\text{a}})}{J(t_{\text{ref}},t_{\text{a}}) - J(t_{\text{bc}},t_{\text{a}})} \approx \frac{\delta(t)}{\delta(t_{\text{ref}})} \tag{8}$$

This formula was verified using the finite element solution of the deflection history of the KB Bridge; see [26, 31]. A detailed analysis of long-term deflection and shortening trends for prestressed box girders is given in [32].

A number of complications arise from the data when optimizing the model. Measurement errors in creep and shrinkage data are a serious obstacle which are discussed in abundance in this and the partner papers. An even more serious complication is introduced by the

enormous amount of creep data that is lacking a reliable documentation of the elastic or instantaneous deformation which anchors the total compliance function. The well established relationship between compliance function $J(28+\Delta,28)$ and standard 28-day modulus E_{28} provides a solution to this problem, if this approximation is introduced as an additional element in the formulation of the objective function. Utilizing all the data with well documented modulus measurements it can be shown that $\Delta = 0.001$ day $\approx$1.5 min provides the best agreement, see [3]. Thus, the contribution of the creep laboratory database to the objective function, considering a relative weight $W_{E_{28}} = 1$ for this additional constraint, reads:

$$\omega_J = \sqrt{\sum_{i=1}^{N_y} W_i \left(\sum_{j=1}^{m} w_{ij} \sum_{k=1}^{n} \left(\frac{y_{ijk} - \hat{y}_{ijk}}{\bar{y_i}} \right)^2 + W_{E_{28}} \frac{E_{28} - J(28+\Delta,28)}{E_{28}} \right)} \tag{9}$$

Different models may best be evaluated and mutually compared by means of the overall C.o.V., calculated as the RMSE of N data points relative to data mean. To avoid statistical bias, it should correctly be written as:

$$C.o.V. = \frac{1}{\sum_{k=1}^{N} w_k y_k} \times \sqrt{\frac{\sum_{k=1}^{N} w_k}{\left(\sum_{k=1}^{N} w_k\right)^2 - \sum_{k=1}^{N} w_k^2} \sum_{k=1}^{N} w_k (y_k - \hat{y}_k)^2} \tag{10}$$

Another statistical indicator that has also been computed is the coefficient of determination (or correlation). Although it does not measure the quality of data fit, it quantifies how closely the data trend is reproduced;

$$R^2 = 1 - \frac{\sum_{k=1}^{N} w_k (y_k - \hat{y}_k)^2}{\sum_{k=1}^{N} w_k (y_k - \bar{y}_k)^2} \tag{11}$$

7 Optimization algorithm

The development of a multi-decade prediction model can actually be seen as two partly independent optimization problems: (a) the functional optimization, i.e. the search for an analytical relationship of the model parameters to the intrinsic and extrinsic input characteristics, and (b) the actual identification of unknown empirical constants. Both, especially the latter, have been extensively studied in various fields.

Apart from the uncertainties of the functional form of prediction equations, which is only partly known from theory, and of the input quantities and measured data, the main challenge lies in the highly nonlinear nature of the optimization problem. The objective function to be minimized typically lacks convexity, which means that local minima exist. Also, it is not 'smooth', and the minimization problem may be mathematically ill-posed [33]. Nevertheless, finding an absolute minimum is made possible by one of the following algorithms: (1) The trust region algorithm (TRA) [34, 35], (2) the genetic algorithms (GA) [36], (3) the particle algorithms (PA) [37], and (4) artificial neural networks (ANN) [38–40]. Here, these mathematical optimization algorithms are used as readily available tools, as discussed later. ANNs implicitly capture the relationship between the observed input and output quantities by optimizing a set of weights during the training phase. ANNs have been successfully applied for inverse identification problems [41], even in cases where no clear functional relationship between inputs and response could be formulated.

The initial parameter and convergence studies show that the optimization problem is sufficiently stable for standard fast gradient-based optimization algorithms, provided that, for a given empirical relationship between the input quantities and model parameters, a starting point close to the global optimum (capturing the right trend) can be supplied. Thus, a two-step cascaded optimization has been used—an initial optimization by, e.g., GA, followed by TRA. Fully relying on GAs has been found to require too many generations. All the present calculations were carried

out in MATLAB, relying on the respective GA and TRA implementations.

8 Functional optimization

The first optimization problem, the functional optimization, consists in a suitable choice of semi-empirical equations relating the intrinsic material properties, geometric characteristics, and environmental conditions to the parameters of the creep and shrinkage model. These are primarily the horizontal and vertical scaling parameters of the different model components: the horizontal scaling of the drying shrinkage curve by the shrinkage half-time, τ_{sh}, its vertical scaling by the final shrinkage, $\epsilon_{sh\infty}$, similar parameters of the autogenous shrinkage model, τ_{au} and $\epsilon_{au\infty}$, and the vertical scaling parameters of the creep compliance function, q_1 to q_5. From Fig. 5a and c, it can be clearly deduced that while a vertical scaling is sufficient to match any feasible compliance curve for basic creep, the same is true for neither shrinkage nor drying creep. In the latter case, one clearly needs the scalings by both the characteristic half-time and the final asymptotic value.

Since no theoretical model for the effects of concrete composition and strength exists, empirical functions must be used. The simplest is to assume linear functions in the logarithmic scale because what mainly matters are the percentage errors rather than actual errors. Therefore, products of power functions, $\prod x_i^{p_i}$, are systematically introduced. They can capture convex as well as concave functional dependance on input quantities x_i, with exponents p_i as the optimization variables. One advantage is that, in its simplest form, this approach can be transformed to linear programming [10].

A major complication in developing and calibrating creep and shrinkage models lies in the 'ill-posedness' of the mathematical optimization problem due to the similarity of various super-imposed time functions. Particularly for shrinkage, it can be shown that an experimental data set that misses the approach to the final value can be closely matched with many different combinations of parameters τ_{sh} and ϵ_{sh}; see Fig. 5d. The same problems afflict the autogenous shrinkage and drying creep components.

However, the sensitivities to certain input quantities and model components change over time. Thus, a sensitivity study of the time function can reveal how to separate model components. Figures 6 and 7, for example, illustrate this fact for the case of a compliance function with a 30 % model error in any one of the model parameters; Fig. 6f shows the relative contribution of the individual compliance terms as part of the total error, for every time after load application $t - t'$. From this plot it is obvious that, with increasing time, the significance of the q_2-term decreases while q_4 and q_5 gain in importance, especially for multi-decade durations for which hardly any laboratory data exist.

Apart from the functional form, an equally important step is the selection of contributing input quantities. Two optimization techniques can be independently or jointly used to uncover relevant inputs: correlation studies and 'sub-setting'. In the former case, potential correlations between model parameters and input quantities (material, environmental, geometrical) are investigated. Inputs with a clear positive or negative correlation are subsequently considered essential. This can serve for the pre-selection of the most important inputs. However, traditional correlation analysis can uncover only linear dependencies and might return inconclusive answers, even if there is a very strong nonlinear dependence.

'Sub-setting' is an extension of the concept introduced with hyper-box weighting. It can be used either for positively identifying important input quantities or for helping to exclude questionable ones. 'Sub-setting' requires the formulation of an n-dimensional hyperspace, in which every input represents a dimension. Subsequently the hyperspace is divided along one dimension into equally sized segments, after which independent correlation studies and optimizations are performed for every 'hyper-slice.' The results for each analysis should be similar, provided the quantity that was chosen for 'sub-setting' is independent. Conversely, if there are distinct differences and trends between the results of individual optimization runs, then the parameter is important and typically involved nonlinearly.

A regression analysis along the dimension of the sub-setted parameter can further reveal the type of functional relationship. The 'sub-setting' is best applied as the first step during the actual parameter identification and model calibration. The input quantities that are assumed to be non-essential are chosen for 'sub-setting' and can be eliminated from further investigation if the results of the analysis confirm the initial assumptions.

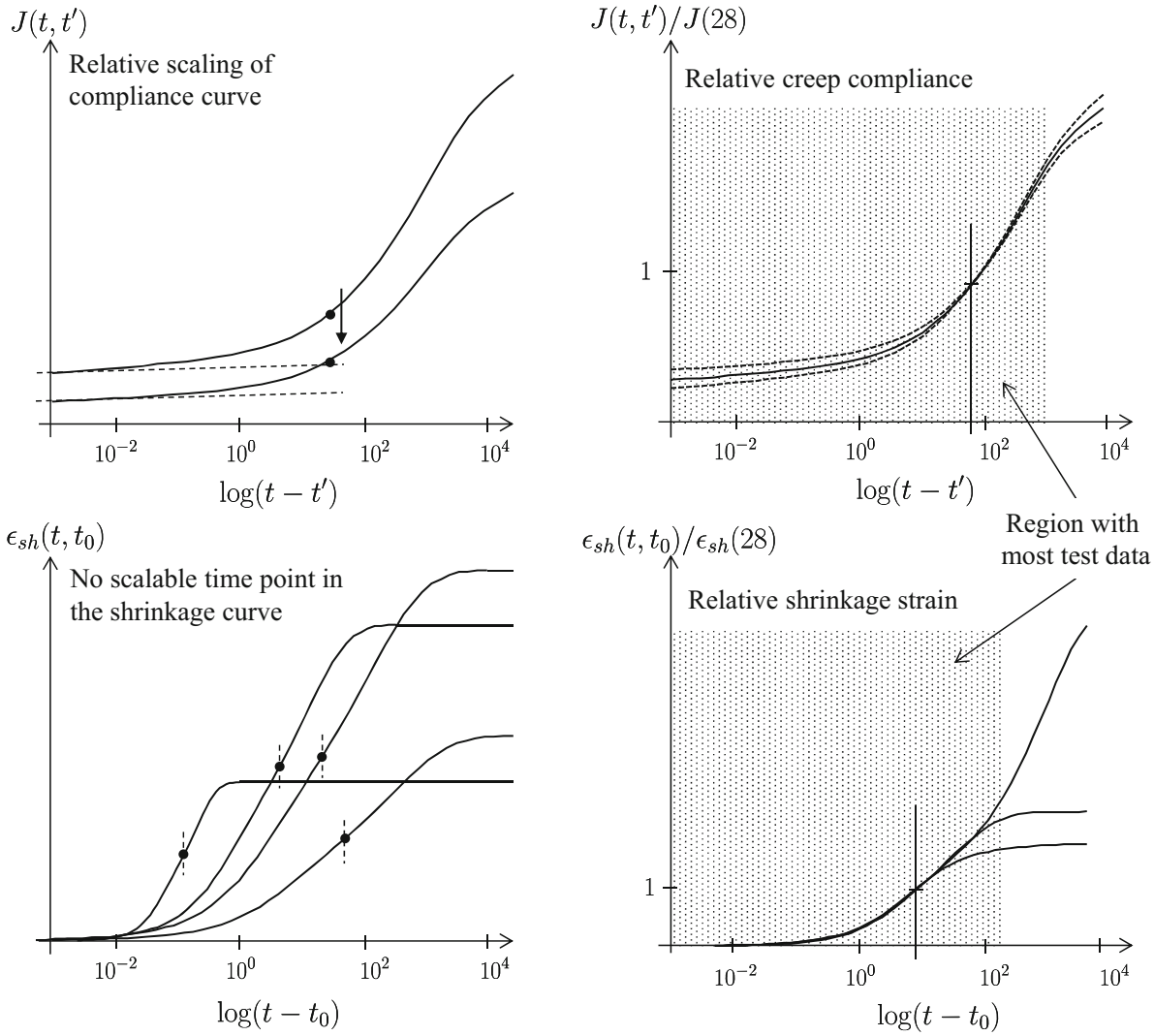

Fig. 5 Relative versus absolute shape

The second important strategy is a 'sub-setting' for the dominant inputs that might influence many model parameters (such as the cement type), and thus temporarily remove it from the optimization problem. This way parameters of lesser sensitivity can be stably identified, using all the information contained within the main 'hyper-slices'.

9 Parameter identification

For a given functional form of the predictor equations (selected on the basis of functional optimization in the first stage), the unknown empirical factors and exponents for the vertical and horizontal scaling parameters must be identified in the second stage of the optimization problem. Typically, functional optimization and parameter identification have to be run alternatingly, until one finds a satisfactory functional form involving as few intrinsic input variables as possible, though not fewer. Based on the interdependence structure of the four model components, the autogenous shrinkage parameters are optimized first, followed by the parameters of drying shrinkage, basic creep and drying creep.

The actual identification of the unknown constants in the predictor equations suffers from the same problems as already discussed, namely the ill-

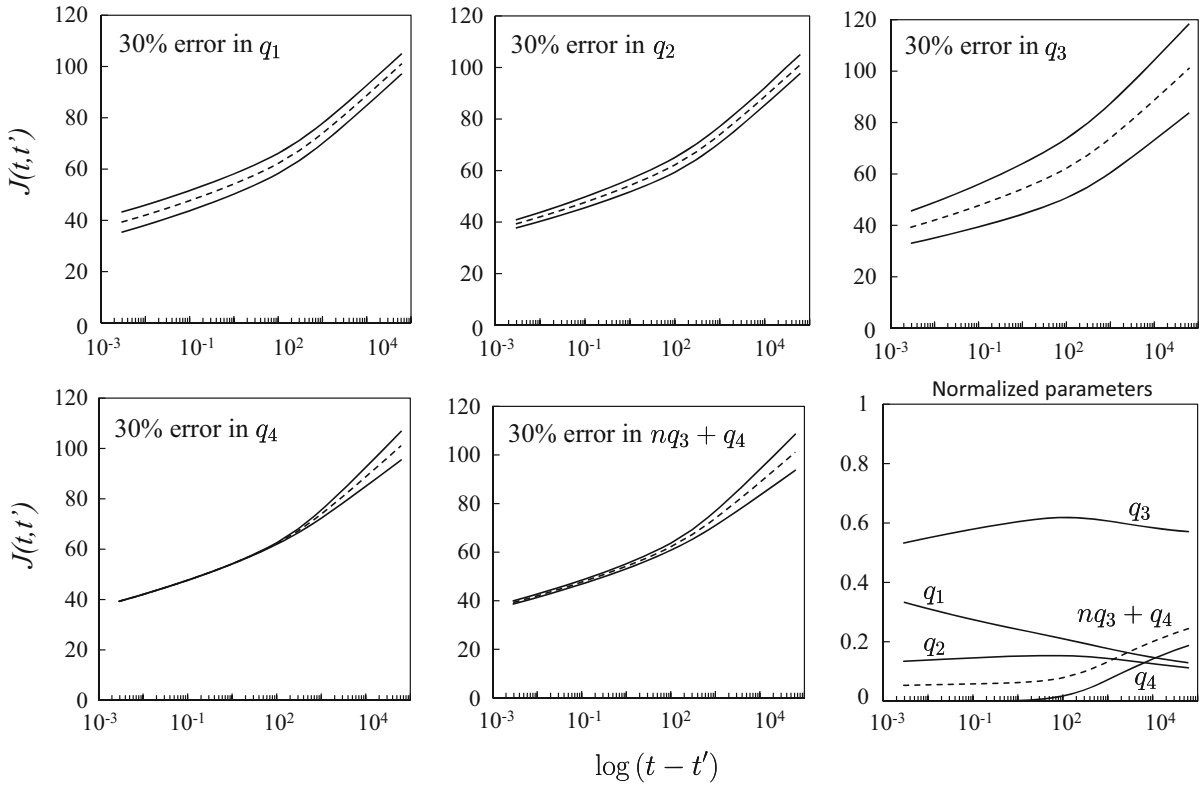

Fig. 6 Predictions of an average test considering the individual sensitivities of different parts of the creep model for basic creep. A 30 % error in the individual q parameters causes a different amount of over- and under-prediction throughout the decades that the test was performed. The last figure illustrates the normalized contribution of each

posedness due to superimposed S-shaped time functions and the uncertainties in the experimental data. These uncertainties are different from the general uncertainties in the mix and composition parameters, the environmental conditions and the precision of the testing procedure; see [1]. They represent measurement errors in the compliance and shrinkage strains, as well as time errors.

If the compliance curve must be reconstructed from the creep coefficient and if the initial elastic deformation either went unreported or was reported without specifying the rate or duration of load application, the entire compliance curve has an unknown vertical shift. If the first reading of shrinkage is delayed for tens of minutes, rather than a few seconds, after the moment of drying exposure, serious errors inevitably occur, shifting vertically the entire shrinkage and drying creep curves. Aside from the random scatter, additional errors are often caused by an improper test setup, by placing the measurement gauge only on the surface of a drying specimen not far enough from the ends, by not sealing the specimen ends, by not using long enough test cylinders, etc. (for details, see [1, 42]). These uncertainties may severely impair the parameter identification [43].

Several strategies to identify and compensate measurement errors in the course of optimization have been investigated. One well established fact, already mentioned, is that ϵ_{sh} initially evolves as $\sqrt{t - t_0}$ where $(t - t_0)$ is the drying duration [44]. Another [4, 45] is that $J(t, t')$ initially evolves approximately as $(t - t')^n$ with $n \in [0.08, 0.12]$. This knowledge can be exploited to identify the time and strain shift errors $(\Delta t, \Delta\epsilon)$ in the reported data—by maximizing the coefficient of determination, R^2, of a linear regression for the initial data range, provided that data for short enough times are available; see Fig. 8. Unfortunately, in practice there are limitations; foremost, most data sets in the laboratory database do not contain enough measurement points in the early part of the curve to allow determining the curve shifts.

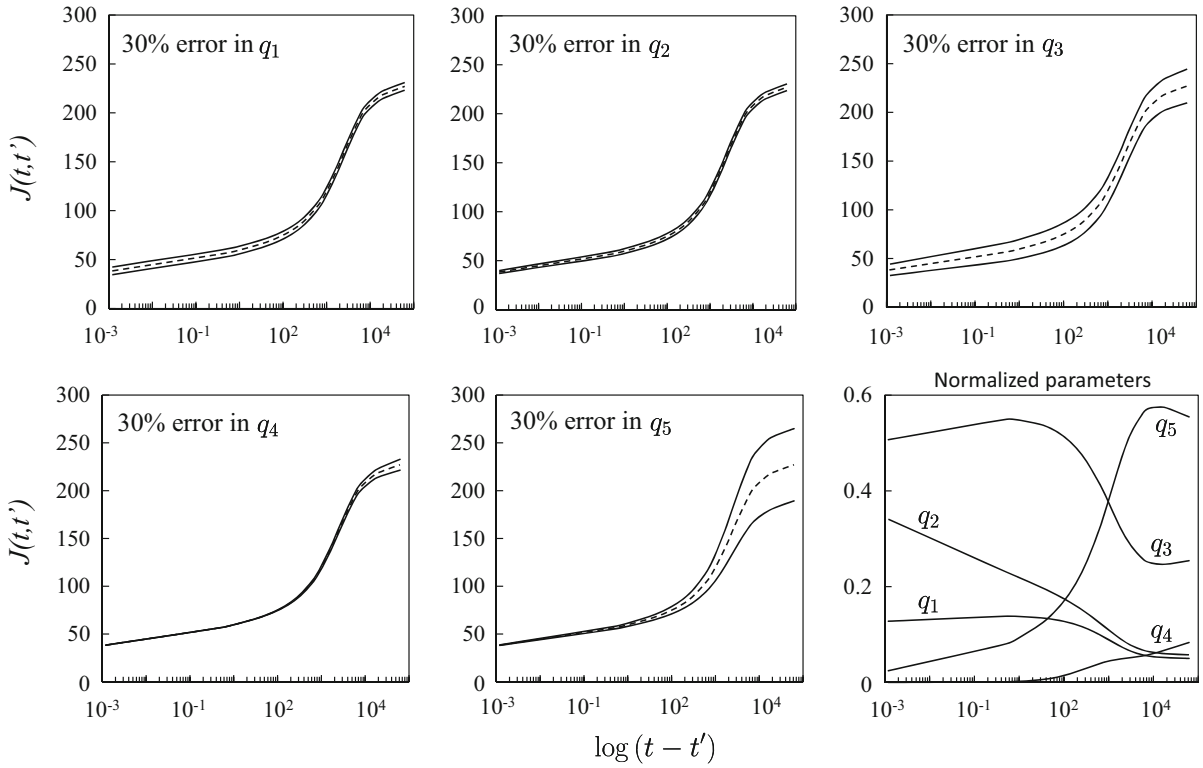

Fig. 7 Predictions of an average test considering the individual sensitivities of different parts of the total creep model. A 30 % error in the individual q parameters causes a different amount of over and under prediction throughout the decades that the test was performed. The last figure illustrates the normalized contribution of each

An alternative approach may involve the normalization of measurement data to the value at a given reference time. This practice, however, leads to reduced sensitivities of the relative time series, as illustrated in Fig. 5, on top of the instabilities arising in a fully automatized normalization of scattered data (the data at or near the chosen reference time might be unavailable or adulterated by scatter). Ultimately, an iterative procedure for identifying the measurement errors, formulated as a sub-optimization problem, proved to provide a robust and unrestrictedly applicable approach; see Fig. 9.

After pre-processing the raw data (basic error checks, application of default input properties, etc.), the bounded error identification and model parameter identification are run alternatingly. Computations showed that 3–5 iterations suffice for convergence, provided that the initial parameter set for the model optimization is close enough to the global optimum. Since Model B4 (as well as B3) defines the initial shrinkage and creep to evolve as $\sqrt{t-t_0}$ and $(t-t')^n$, respectively (with $n = 0.1$), this approach not only reduces errors in conformity to the theory but also enhances the applicability range.

The iterative procedure to identify initial errors in time and strain shifts was used to calibrate both the creep and shrinkage models (path A of Fig. 9). While this procedure could be consistently used for shrinkage, its application was limited to later phases of the creep model optimization. In particular, the optimization of the initial vertical offset in the instantaneous creep compliance term required the use of reliable actual (rather than relative) data (see subsequent section on the creep specifics), which allowed no shift (path B).

The subsequent steps of the optimization utilized the iterative error optimization strategy, which was bounded in order to avoid numerical instabilities or impossible shifts. The bounds were ± 1 day for Δt, $\pm 20\mu$-strains for $\Delta\epsilon_{sh}$, and $\pm 200 \times 10^{-6}/$ MPa for creep. Negative ages or function values were, of course, excluded. A

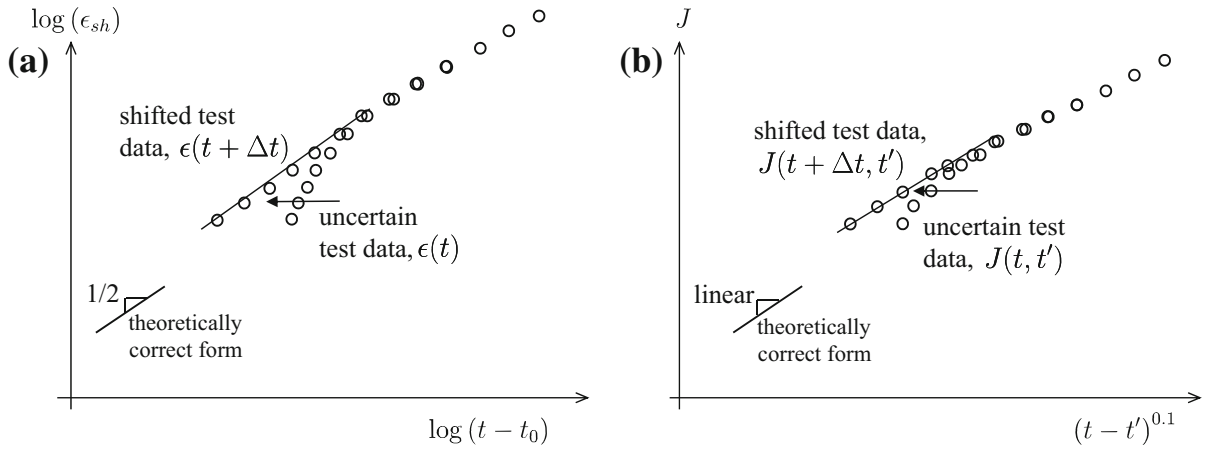

Fig. 8 Direct identification of time shift error; find Δt such that the linearity of the initial test data be maximized

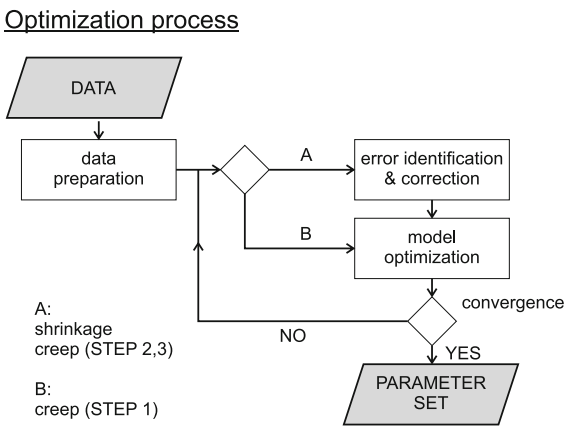

Fig. 9 Iterative strategy for the identification of measurement errors $\Delta t, \Delta J, \Delta\epsilon$ in creep and shrinkage

relatively large range for creep is required to deal with the initial compliance measurements, which are frequently missing or severely distorted.

To verify the chosen approach to dealing with measurement errors, the model was repeatedly validated directly against the raw data. Also, all the statistical data analyses were performed directly on the raw data as well as in combination with the aforementioned error optimization procedure. Figure 10 summarizes the obtained shift distribution for the final, purely predictive, runs of the creep and shrinkage model. The slim outermost columns give the number of curves that were hitting the bounds prescribed for the error identification.

While both creep and shrinkage reveal the need to correct time shift errors, the significance of errors in the function value is more pronounced for shrinkage. Almost none of the creep data sets hit the prescribed bounds. For the majority of data, the shift $\Delta J \approx 0$. The obtained distribution shows a positive skew, which suggests underestimation of the compliance due to probable omission of measurements for very short creep durations.

The individual optimization strategies for creep and shrinkage, which are discussed in the next two sections, are formulated in such a way as to allow an evolving model instead of a brute-force simultaneous optimization of all variables. The model complexity and the number of variables increase in several steps, in which the uncertainty in assumptions and the sensitivity to input quantities decrease. The alternative approach—optimization of the full model in one step—would have been highly ill-conditioned, especially in view of the measurement errors and uncertainties in the raw data.

A further requirement is the careful selection and iterative refinement of the optimization bounds, which have to be chosen by expert judgement. Moreover, since the optimization procedure should ultimately arrive at a 'nice' form of the derived empirical predictor equations, the parameter space is, by design,

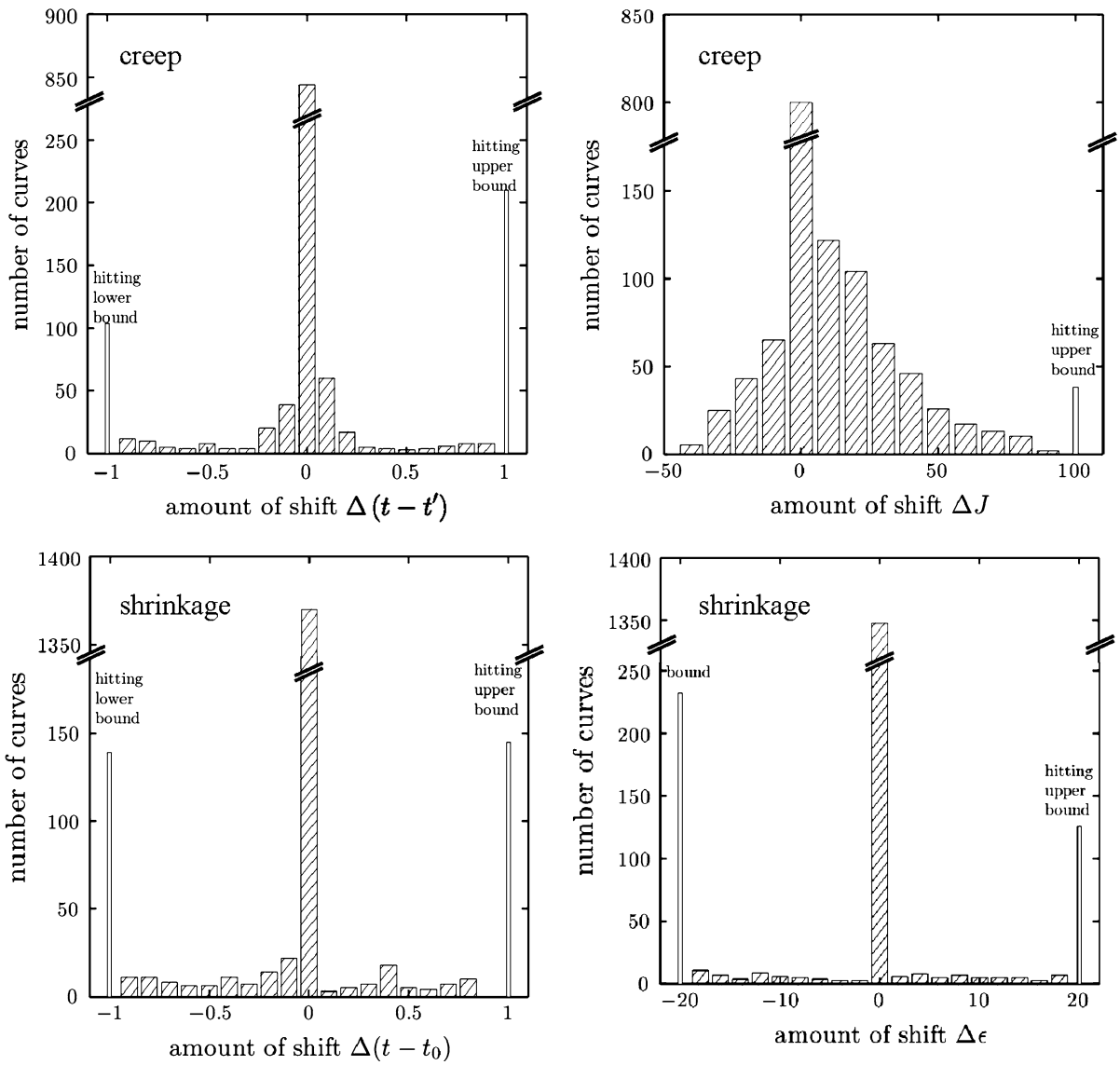

Fig. 10 Distribution of optimized shifts for the creep model (*top*) and shrinkage model (*bottom*); shifts in measurement value (*left*) and in time (*right*)

discontinuous. It follows that, depending on the chosen precision, the optimization either is performed on stepwise constant parameters, which introduces convergence problems, or involves final rounding of the optimization variables. This had to be done iteratively, with decreasing sensitivity of the optimization variable, to allow for the required slight redistribution of information.

Each step of the calibration follows a similar strategy as already discussed for the model development and functional optimization. The initial parameter space is subset based on one or several important input quantities that are not yet considered in the model (e.g., cement type and admixtures). Consequently, the data in each subset have consistent properties, which are fitted more easily. Information on the dependence of the model on the subset parameters (or a lack thereof) is obtained automatically.

A bigger challenge is how to extract clear and physically meaningful trends. To capture them, correlation analyses and sensitivity studies must be repeatedly conducted in each optimization step. High correlations

between input quantities and optimization variables are a clear indication of redundancy, which must be addressed.

Maximum variance sub-setting can help to confirm trends. This technique involves subdividing the available data by random assignment into several subsets in such a way that, in each subset, the variance of the non-subsetted input quantities be maximized. This creates multiple independent and representative data sets for optimization with sufficient and representative sensitivity to the input quantities.

The sensitivity of the overall optimization to single input quantities (and its change in time) is best characterized by the Jacobian, $\partial F/\partial \mathbf{X}$, of the objective function F with regard to the optimization variables $\mathbf{X}$. Figure 11 shows a plot of F in a normalized form, in which the regimes with different parameter sensitivities are identified (see the creep optimization section for further discussion). Note that, due to the change of parameter sensitivity with time, various weighting schemes, e.g., putting more weight on certain time ranges, might be helpful, depending on the parameter to be identified. However, within this investigation, a consistent weighting scheme as described above was applied.

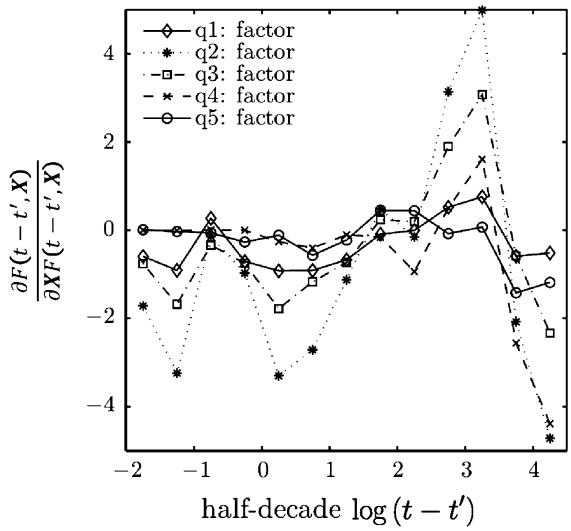

Fig. 11 Evolution of the Jacobian over time as calculated in half-decades in logarithmic time. The q_2 value is the most sensitive over the full time span; the q_3 parameter and in particular q_4 gain in importance for longer time spans

10 Optimization strategy for shrinkage

Among the four components of Model B4, calibration must begin with autogenous shrinkage and drying shrinkage, following the flow chart in Fig. 12. Of these two, the autogenous shrinkage, which is independent of the environmental humidity, is calibrated first (step 1). It already provides some level of prediction for the identification of the drying shrinkage model parameters in step 2, which also depends on the chosen interaction rule. For autogenous shrinkage during simultaneous external drying, water diffusion analysis would be most realistic. Unfortunately, such an approach is hardly feasible for a model giving the average response of the cross section. Therefore, the autogenous and drying shrinkage strains are simply assumed to be additive.

The fact that the autogenous and drying shrinkage curves are somewhat similar in shape introduces in the optimization some degree of ill-posedness, since different combinations of these curves can yield almost he same coefficient of variation. This, and a high number of unknown influences, required the development of a multi-stage multi-step optimization strategy in which the number of unknown model parameters increased greatly above the initial value of $N = 7$ for the most basic formulation. In the first phase, differences in cement type, admixture content, temperature and aggregate type have been ignored, which yields this first model. Where required, default values are taken from the literature; e.g., for the activation energies of temperature effects. This initial step is most important, as the final form of the model is decided and the most relevant trends and correlations to the input quantities are revealed.

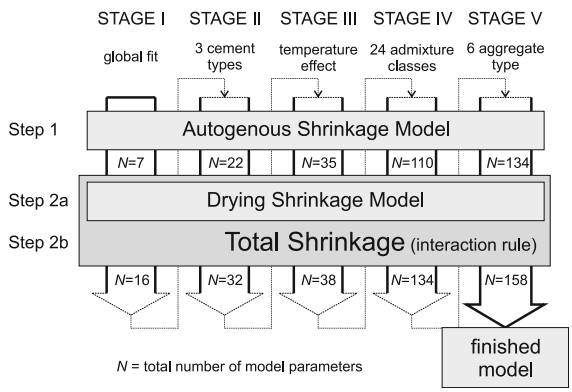

Fig. 12 Optimization strategy for shrinkage

After a satisfactory fit is obtained, further model refinements are introduced in the sequence of their importance and likelihood to improve the quality of fit. Subsequent refinements include the consideration of the cement type (which considerably improved the predictive capability), admixtures, aggregate type, and the temperature effect; in more detail, see [2, 3]. Note that while not all of the N parameters could remain open simultaneously, the most sensitive ones were refitted at each model refinement (Fig. 12).

The analysis of several hundred journal articles, reports and theses reporting shrinkage tests [1] revealed that while measurements of autogenous shrinkage begin at the time of set of the concrete mix, most of the measurements of drying shrinkage begin only after exposure to the environment, typically at 7 to 28 days of age. However, when the modern concrete specimens for drying shrinkage undergo curing, a significant autogenous shrinkage is already under way and continues during drying until the pore humidity drops below the self-desiccation humidity, which occurs in the specimen core with great delay. Since the autogenous shrinkage before the exposure to drying usually went unrecorded, the reported total shrinkage, i.e. drying plus partial autogenous, underestimates the true amount.

For the model B4 calibration, the following corrections are applied to data. As sketched in Fig. 13, the observed response $\epsilon_{\mathrm{sh,total}}(t-t_0)$, which is supposed to represent the sum of autogenous shrinkage ϵ_{au} and drying shrinkage ϵ_{sh}, is missing the autogenous shrinkage part, $\epsilon_{\mathrm{au}}(t=t_0)$, that occurred prior to the exposure of the specimen to the environment at time t_0. Therefore, during fitting, this real situation is imitated by subtracting the predicted value $\hat{\epsilon}_{\mathrm{au}}(t_0)$ from the total shrinkage prediction.

11 Optimization strategy for creep

The optimization strategy for creep follows similar principles, in five stages of increasing model complexity and number N of optimized variables; Fig. 14. Instead of the autogenous and drying shrinkage components, it is now the asymptotic instantaneous compliance (q_1), basic creep compliance and drying creep compliance terms that need to be separated. While the time functions for ageing and non-ageing viscoelastic creep ($q_2 - q_3$) are theoretically coupled, the flow term (q_4) and the drying creep compliance term (q_5) are uncoupled. The lack of multi-decade creep data is compensated by considering the data on relative bridge deflections, converted to relative compliance, as already explained.

Step 1 is the most sensitive aspect of the optimization problem, because the instantaneous compliance anchors the entire multi-decade prediction model. Due to widespread problems with measuring or reporting this compliance, only the most reliable data are used. A compliance shift ΔJ would conflict with the optimization target q_1, which represents the asymptotic ordinate (or intercept). The approximate existence of a common left-side asymptotic ordinate of compliance curves for different ages at loading was proven in Fig. 6 of [4] and is further supported in the next article [3].

Step 2 optimizes the basic creep compliance parameters, first using only the short-term data (with higher sensitivity to q_2 and q_3), and then the long-term data to calibrate q_4 for the flow term.

Step 3 optimizes q_5 to fit the drying creep data jointly with the basic creep data. The calibration of q_5 is combined with a recalibration of q_2, q_3, q_4 within narrow bounds. The final step 4 recalibrates q_3 and q_4 jointly with the converted relative bridge compliance information.

The available bridge data provide no insight into the influence of temperature on creep and its dependence on the environmental humidity. It must be noted that the converted bridge data are biased towards the typical average composition and strength values which, for a lack of information, had to be assumed for the concretes

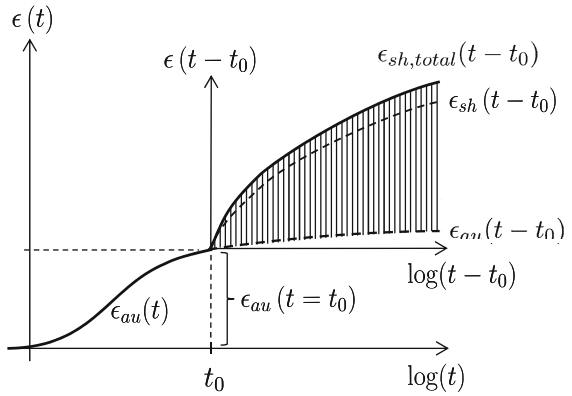

Fig. 13 Dealing with missing autogenous shrinkage contribution

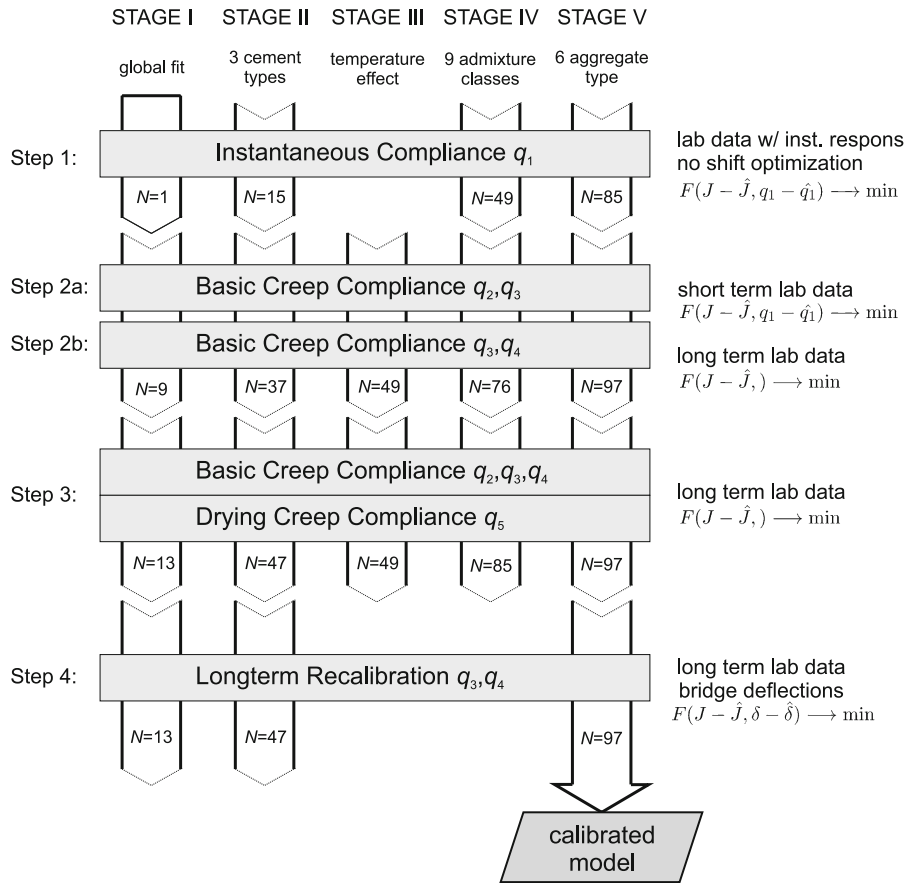

Fig. 14 Optimization strategy for creep

of the 69 bridges in the database. These typical properties were assumed as $w/c = 0.5$, $c = 400$ kg/m^3 (25 lb/ft^3) and mean strength = 39 MPa. Bias towards the assumed composition parameters and environmental conditions in the calibration of model B4 is avoided by opening solely the scaling factors associated with q_3 and q_4 and simultaneously keeping all the composition related terms locked during step 4.

The described procedure (steps 1–4) is repeated for all the five stages of the optimization, in the course of which the parameters were seen to stabilize.

The evolution of the sensitivity of the model to the individual compliance terms is plotted in Fig. 11 in terms of the normalized Jacobian of the objective function $\partial F/(F\partial \mathbf{X})$.

12 Comparisons of various models

Based on the statistical indicators, the new Model B4 with composition and strength effects, as well as the simplified Model B4s which considers only the strength effect, are analyzed to answer three questions: (1) How closely can the models capture the shape of the observed individual test curves? (2) How closely can the models reproduce the characteristic trends with respect to temperature, environmental humidity and specimen size? (3) Can the models represent the entirety of the data collection in a statistically acceptable sense?

To answer these questions, both models are applied to: (a) selected sets of individual curves with sufficient data in the initial and final phases; (b) selected subsets of experiments of the same concrete in which only specified extrinsic parameters were varied; and (c) the entirety of the available data, in a predictive way. Both models are compared to the RILEM recommendation B3 [4], the *fib* model code 1990–1999 [46], its successor—the 2010 *fib* model code [13] as updated in 2012, the Gardner–Lockman model 2000 [47], and the ACI 1992 model [14] (which is virtually identical to the ACI 1971 model [48]).

13 Verification of functional form by fitting individual creep and shrinkage curves

As already emphasized, the first check of any prediction model, important especially for fully empirical models, is that its functions must have a mathematical form that is, after optimization of its parameters, able to fit closely the individual test curves measured on one and the same concrete, so that the enormous obfuscating scatter due to variations of concrete composition be avoided. If the functional form of a model cannot provide a very close fit of such individual measured curves, one need not even bother to optimize the model to the complete database.

A set of test curves for the same concrete should include the evolution of strain in time and preferably also the effects of the age at loading, environmental humidity and member thickness. One must select a few trustworthy data sets with sufficient data in the initial short-time range as well as the terminal time range on approach to a (rising) asymptote. Naturally, the prediction for creep must be scaled vertically, and for shrinkage both vertically and horizontally, to account for the specific characteristics of the given concrete which typically deviate from the calibrated average.

Simple weighted least-square fitting suffices for this essential first check, although here a full weighted optimization as discussed above is performed. Unfortunately, despite its simplicity, this basic check has often been omitted in the evaluation of most models.

14 Subsets of individual data sets

In the next phase, all models (B3, B4, MC, GL, ACI) are refitted to a selection of approximately 20 shrinkage tests and 40 creep tests that yielded trustworthy measurements of sufficient range, including early measurements, and showed a variation in one extrinsic input quantity only (e.g. time of loading, temperature, relative environmental humidity). The fitting variables are chosen according to the respective model formulations with the sole purpose of recalibrating the individual models to the specific properties of the test data at hand.

The updated parameters are limited to vertical scaling of the final shrinkage value (autogenous and drying) $\in$ [0.3,3.0], horizontal scaling of the characteristic halftimes $\in$ [0.2, 5.0], and separate vertical scaling of each compliance term. To prevent measurement errors from interfering, this analysis is performed once using the actual measurement values, once using the values normalized to the 7-day prediction, and once using the values normalized to the 28-day prediction (a detailed discussion of this comparison is presented in the subsequent two companion articles on creep [3] and shrinkage [2]). As already mentioned, a comparison of the fitted relative data is of little value, especially for shrinkage.

Detailed analyses of the capability of the model to reproduce the behavior for each single concrete and the major parameter trends provide invaluable insights that cannot be achieved by overall statistical analysis.

15 Global statistics

The second part of any model validation and comparison should consist of a comprehensive statistical analysis of all available data—in this case the entirety of the new NU-RILEM creep and shrinkage databases [1]. The new models B4 and B4s are directly compared to the major prediction models endorsed by various engineering societies using the C.o.V. (normalized root-mean-square error) as the statistical indicator. Contrary to the previous analysis, all models are applied in a purely predictive way to a subset of the full database that complies with the requirements of each model (available input quantities and their range). Potential errors in measurement and time are ignored in stage 1 of the comparison and are dealt with in stage 2 as defined for the identification procedure. The shift identification was applied non-discriminately for all models.

Both approaches led to comparable results, i.e., the same ranking of the models as discussed in detail in the companion articles for the creep [3] and shrinkage [2] analyses. Figures 15 and 16 show the amount of available data points and the quality of fit respectively.

16 Uncertainty quantification

For many practical applications it does not suffice to know that a model can provide good average predictions for one specific concrete and for all concretes. One

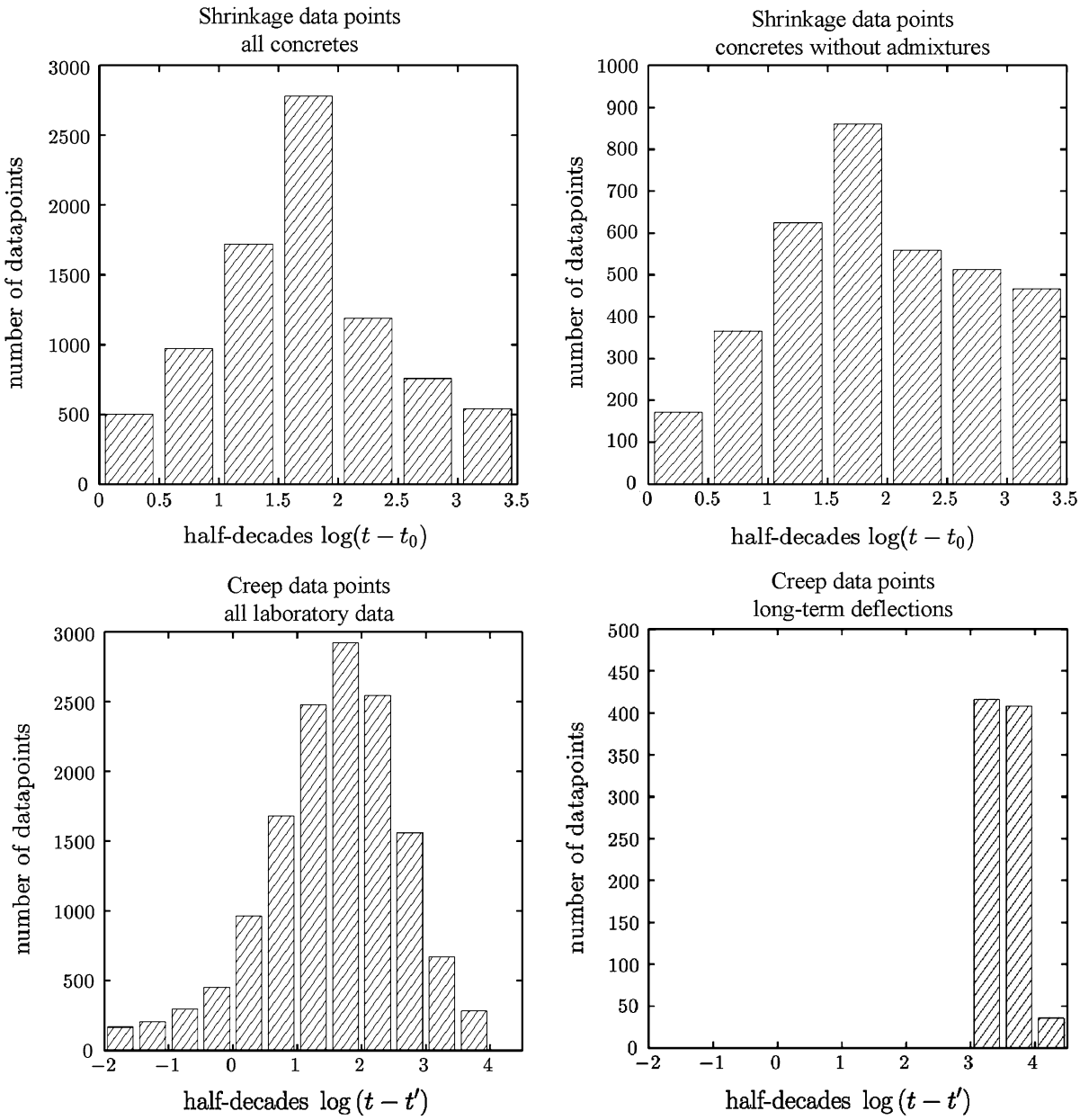

Fig. 15 Number of data points available for calibration

must also quantify the uncertainty of the prediction. Two types can be distinguished—the aleatory and epistemic uncertainties [49]. While the former represents an inherent statistical uncertainty that cannot be avoided, the latter covers the model uncertainty, i.e., the systemic uncertainty that, in principle, could be known. The typical goal of uncertainty quantification is a reduction of the epistemic uncertainty and transformation to aleatoric uncertainty. Assuming a perfect model, aleatoric uncertainties can be quantified in a straightforward approach, e.g., by Monte Carlo sampling.

To quantify the uncertainty of models B4 and B4s, we introduce an inverse uncertainty assessment in which all the available data sets (creep and shrinkage curves) are refitted individually by the prediction model. For shrinkage, parameters $\tau_{\mathrm{sh}}, \epsilon_{\mathrm{sh},\infty}, \tau_{\mathrm{au}}, \epsilon_{\mathrm{au},\infty}$ are considered uncertain. For creep, these are $q_1, \ldots, q_5$. This assessment leads to the distributions ψ_i of the necessary

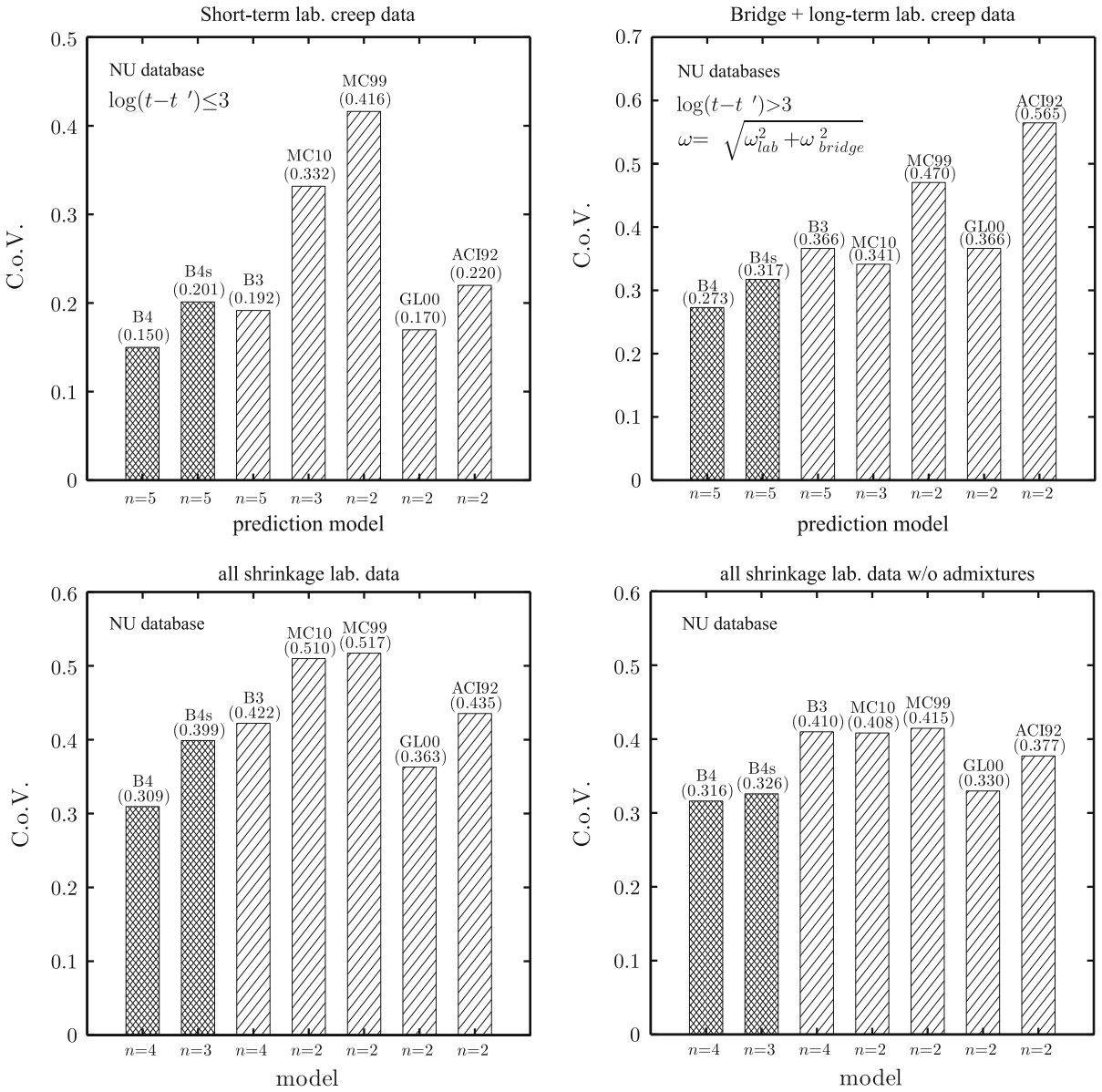

Fig. 16 Quality of fit of the new B4 and B4s models as compared to existing models using the C.o.V. as the quantifier

correction factors of the uncertain model parameters p_i which are required to match the specific data at hand with the average prediction. The correction factors ψ_i primarily capture the intrinsic material uncertainty, biased by random errors associated with the testing.

To ensure numerical stability, bounds are imposed on the correction factors: $\psi_i \in [0.3, 3.0]$. After removing the bounded samples, the resulting empirical distributions could be fitted closely with log-normal distributions. This distribution type is justified by the central limit theorem according to which the product of n independent random variables approaches a log-normal distribution and gives always non-negative scaling factors. A detailed discussion of the resulting uncertainty factors for the 5 and 95 % cutoffs is given in [2] and [3].

17 Bayesian updating approach

For a particular structure, the creep and shrinkage predictions may be updated if the concrete mix is known, if observations on this or similar structures are available, or if multi-decade data on a similar concrete exist. For this purpose, the framework of Bayesian updating [50, 51] is most suitable.

Assume that a state S occurs with a probability $P(S)$ before any observation M is taken. The probability $P(S)$ is, in this context, called the prior probability. The posterior probability $P(S|M)$ given an observation, M, can be calculated according to Bayes' rule [50, 51] by:

$$P(S|M) = \frac{P(M|S)P(S)}{P(M)} \tag{12}$$

Here $P(M|S)$ is the probability of observing M given a certain state S. It is usually referred to as 'likelihood'. $P(M)$ is sometimes termed 'model evidence' or 'marginal likelihood' and is independent of the hypothesis or state investigated.

According to the concept of Bayesian updating, the data on multi-decade deflections δ of many bridges are used as the evidence to improve the creep prediction model calibrated by laboratory creep data most of which are of much shorter durations. Recall the functional form of B4: The creep compliance function is given by a sum of the instantaneous compliance with parameter q_1, aging viscoelastic compliance with q_2, non-ageing viscoelastic compliance with q_3, ageing flow compliance with q_4 and drying creep compliance with q_5. Fitting of the laboratory database, mostly short-time, can provide the uncertainty quantification of the prior distributions of $q_1, \ldots, q_5$. Since the terminal slope is determined only by q_3 and q_4, a simplified updating of solely these two parameters is sufficient. In that case, the informed posterior distribution may be written as $f_Q(q_3, q_4|U = \delta)$, while the prior $f_Q(q_3, q_4)$ is taken from the uncertainty quantification. Hence,

$$f_Q(q_3, q_4|U = \delta) = \frac{f_U(\delta|Q = q_3, q_4) f_Q(q_3, q_4)}{\int_{-\infty}^{\infty} f_U(\delta|Q = \bar{q_3}, \bar{q_4}) f_Q(\bar{q_3}, \bar{q_4}) d(\bar{q_3}, \bar{q_4})} \tag{13}$$

The same framework can naturally be applied to improving the model calibration during construction, provided that short-term data on the concrete used are available or that data from the direct or indirect monitoring of deflections exist [32, 52, 53].

18 Concluding comments

Why have simplistic empirical models for creep and shrinkage prediction survived for such a long time? The likely reason is that, at the time of design, the potential creep problems are normally at least 20 years in the future. After twenty or more years, problems can easily be blamed on a variety of extraneous effects such as poor quality control, excessive loads, corrosion damage, etc., rather than on the error of a standardized formula endorsed long ago by a large committee of structural engineers. The problems due to shrinkage get manifested earlier but they, too, can easily be blamed on various extraneous effects.

Recently, after demonstrations of the widespread loss of long-term durability, the enormous costs of shortened lifetimes are being recognized. It has become clear that the problems of creep and shrinkage in concrete structures deserve keen attention, which includes careful optimizations and adaptations of the prediction models. Unfortunately, these problems are also complex. The necessary complexity inevitably calls for a sophisticated procedure of model optimization and uncertainty quantification, which has been the objective of this article.

The previous models adopted by various engineering societies received only limited calibration by elementary procedures. Their calibration lacked many of the aspects of modern optimization and, in particular, omitted several crucial steps. These steps include:

1. Selecting for optimization a model form that has physical and theoretical support;
2. Checking whether this form can closely fit the available individual creep and shrinkage curves that have been obtained for one and the same concrete and have a time range stretching from short enough to long enough times;
3. Compensating for the limited time range of laboratory databases by taking into account multi-decade observations on structures;
4. Counteracting various kinds of bias and arbitrary data sampling in the available experimental database; and

5. Dealing properly with potential measurements errors and incomplete experimental data.

These steps are inevitable for obtaining a model that could be trusted for multi-decade extrapolations—even predictions for more than 100 years, which would be in line with the current lifetime requirements for large and expensive structures. Particularly important for multi-decade prediction and extrapolation of short-time data for a given concrete is step 2, because it avoids contamination by the huge obfuscating scatter due to the effects of different concrete compositions.

Fortunately, we live in a computer age in which the use of realistic, though complex, mathematical models and procedures is feasible and easy, in fact as easy as the use of simplistic models.

Acknowledgments Financial support from the U.S. Department of Transportation, provided through Grant 20778 from the Infrastructure Technology Institute of Northwestern University, is gratefully appreciated. The first author thanks the Austrian Science Fund (FWF) for additional support in the form of the Erwin–Schrödinger Scholarship J3619-N13. Partial financial support by the Austrian Federal Ministry of Economy, Family and Youth and the National Foundation for Research, Technology and Development is gratefully acknowledged.

References

1. Hubler MH, Wendner R, Bažant ZP (2014) Comprehensive database for concrete creep and shrinkage: analysis and recommendations for testing and recording. ACI (accepted)
2. Hubler MH, Wendner R, Bažant ZP (2014) Statistical justification of Model B4 for drying and autogenous shrinkage of concrete and comparisons to other models. RILEM Materials and Structures (accepted)
3. Wendner R, Hubler MH, Bažant ZP (2014) Statistical justification of Model B4 for multi-decade concrete creep and comparisons to other models using laboratory and bridge databases. RILEM Mater Struct (accepted)
4. Bažant ZP, Baweja S (1995) Creep and shrinkage prediction model for analysis and design of concrete structures: Model B3. RILEM Recommend Mater Struct 28:357–367 (Errata, 29:126)
5. Robert-Nicoud Y et al (2005) System identification through model composition and stochastic search. ASCE J Comput Civ Eng (JCCE) 19(3):239–247
6. Bažant ZP, Prasannan S (1989) Solidification theory for concrete creep: I. Formulation. J Eng Mech ASCE 115(8):1691–1703
7. Bažant ZP, Hauggaard AB, Baweja S, Ulm F-J (1997) Microprestress-solidification theory for concrete creep. I. Aging and drying effects. J Eng Mech ASCE 123(11):1188–1194
8. Bažant ZP, Hauggaard AB, Baweja S (1997) Microprestress-solidification theory for concrete creep. II. Algorithm and verification. J Eng Mech ASCE 123(11):1195–1201
9. Bažant ZP, Carreira D, Walser A (1975) Creep and shrinkage in reactor containment shells. J Struct Div Am Soc Civ Eng 101:2117–2131
10. Bažant ZP, Baweja S (2000) Creep and shrinkage prediction model for analysis and design of concrete structures: Model B3. In: Al-Manaseer A (ed) Adam Neville symposium: creep and shrinkage, structural design effects. American Concrete Institute (ACI SP194), Farmington Hills
11. Bažant ZP, Najjar LJ (1972) Non-linear water diffusion in non-saturated concrete. Mater Struct 5:3–20
12. Bažant ZP, Hubler MH, Wendner R (2014) Model B4 for creep, drying shrinkage and autogenous shrinkage of normal and high-strength concretes with multi-decade applicability. RILEM draft recommendation. TC-242-MDC MultiDecade Creep and Shrinkage of Concrete: Material model and structural analysis. RILEM Materials and Structures (accepted)
13. Fédération Internationale du Béton (2010) Structural concrete: textbook on behaviour, design and performance, vol 51. Int. du Béton, FIB-Féd
14. ACI Committee 209 (2008) Guide for modeling and calculating shrinkage and creep in hardened concrete. ACI Report 209.2R-08, Farmington Hills
15. Strauss A, Hoffmann S, Wendner R, Bergmeister K (2009) Structural assessment and reliability analysis for existing engineering structures, applications for real structures. Struct Infrastruct Eng 5(4):277–286
16. Wendner R, Strauss A, Guggenberger T, Bergmeister K, Teply B (2010) Approach for the assessment of concrete structures subjected to chloride induced deterioration. Beton- und Stahlbetonbau, 105. Heft 12, pp 778–786
17. Strauss A, Wendner R, Bergmeister K, Costa C (2013) Numerically and experimentally based reliability assessment of a concrete bridge subjected to chloride induced deterioration. J Infrastruct Syst 19(2):166–175
18. Müller HS, Hilsdorf HK (1990) Evaluation of the time-dependent behaviour of concrete: summary report on the work of the General Task Force Group No. 199. CEB (Copmité euro-internationale du béton), Lausanne, 201 pp
19. Troxell GE, Raphael JE, Davis RW (1958) Long-time creep and shrinkage tests of plain and reinforced concrete. In: Proceedings of ASTM, vol 58, pp 1101–1120
20. Russel HG, Larson SC (1989) Thirteen years of deformations in water tower place. ACI Struct J 86(2):182–191
21. Burg RG, Orst BW (1994) Engineering properties of commercially available high-strength concretes (including three year data). PCA Research and Development Bulletin RD104T, IL, Portland Cement Association, Skokie, p 58
22. Browne RD, Bamforth PP (1975) The long term creep of Wylfa P. V. concrete for loading ages up to 12 1/2 years. In: 3rd International conference on Structural Mechanics in Reactor Technology (SMiRT-3), paper H1/8, London
23. Hanson JA (1953) A ten-year study of creep properties of concrete. Concrete Lab. Report No. SP-38. US Department of the Interior, Bureau of Reclamation, Denver
24. Harboe EM et al (1958) A comparison of the instantaneous and the sustained modulus of elasticity of concrete, Concr.

Lab. Rep. No. C-354. Division of Engineering Laboratories, US Department of the Interior, Bureau of Reclamation, Denver
25. Brooks JJ (2000) Elasticity, creep and shrinkage of concretes containing admixtures. ACI Journal SP 194, pp 283–360
26. Bažant ZP, Hubler MH, Yu Q (2011) Pervasiveness of excessive segmental bridge deflections: wake-up call for creep. ACI Struct J 108(6):766–774
27. Bažant ZP, Li G-H (2008) Unbiased statistical comparison of creep and shrinkage prediction models. ACI Mater J 105(6):610–621
28. Tofallis C (2008) Least squares percentage regression. J Mod Appl Stat Methods 7:526–534
29. Bažant ZP, Panula L (1978–1979) Practical prediction of time-dependent deformations of concrete. Mater Struct (11) Part I, Shrinkage, pp 307–316. Part II, Basic creep, pp 317–328. Part III, Drying creep, pp 415–424. Part IV, Temperature effect on basic creep, pp 425–434
30. Bažant ZP, Yu Q, Li, G-H (2012) Excessive long-time deflections of collapsed prestressed box girders: II. Numerical analysis and lessons learned. ASCE J Struct Eng 138(6):687–696
31. Yu Q, Bažant ZP, Wendner R (2012) Improved algorithm for efficient and realistic creep: analysis of large creep-sensitive concrete structures. ACI Struct J 109(5):665–676
32. Wendner R, Tong T, Strauss A, Yu Q (2014) A case study on correlations of axial shortening and deflection with concrete creep asymptote in segmentally-erected prestressed box girders. Struct Infrastruct Eng (in press)
33. Buljak V, Cocchetti G, Cornaggia A, Garbowski T, Maier G, Novati G (2013) Mechanical characterizations of materials and structural diagnoses by inverse analyses. Book chapter draft
34. Coleman TF, Li Y (1996) An interior, trust region approach for nonlinear minimization subject to bounds. SIAM J Optim 6:418–445
35. Conn AR, Gould NIM, Toint PL (2000) Trust-region methods. Society for Industrial and Applied Mathematics, Philadelphia
36. Koh CG, Perry MJ (2009) Structural identification and damage detection using genetic algorithms. CRC Press, Boca Raton
37. Lazinica A (ed) (2009) Particle swarm optimization. InTech
38. Waszczyszyn Z (1999) Neural networks in the analysis and design of structures. Springer, Berlin
39. Hagan MT, Demuth HB, Beale MH (1996) Neural network design. PWS Publishing, Boston
40. Haykin S (1998) Neural networks: a comprehensive foundation. Prentice Hall, Upper Saddle River
41. Novak D, Lehky D (2004) Neural network based identification of material model parameters to capture experimental load-deflection curve. Acta Polytech 2006 44(5–6):110–116
42. Acker P, Bažant ZP, Chern JC, Huet C, Wittmann FH (1998) RILEM recommendation on measurement of time-dependent strains of concrete (prepared by Subcomm. 4 of RILEM Committee TC107-CSP). Mater Struct 31(212): 507–512
43. Subcommittee RILEM, 4 (1998) Measurement of time-dependent strains of concrete. Mater Struct 31:507–512
44. Wittmann FH, Bazant ZP, Alou F, Kim JK (1987) Statistics of shrinkage test data. Cem Concr Aggreg ASTM 9(2): 129–153
45. Vandamme M, Zhang Q, Carrier B, Ulm F-J, Pellenq R, van Damme H, Le Roy R, Zuber B, Gartner E, Termkhajornkit P (2013) Creep properties of cementitious materials from indentation testing: signification, influence of relative humidity, and analogy between C-S-H and clays. Keynote presentation. CONCREEP 9, Boston
46. FIB (1999) Structural concrete: textbook on behaviour, design and performance, updated knowledge of the CEB/FIP Model Code 1990. Bulletin No. 2, Fédération internationale du béton (FIB), Lausanne 1, pp 35–52
47. Gardner NJ, Lockman MJ (2001) Design provisions for drying shrinkage and creep of normal strength. ACI Mater J 98(2):159–167
48. Committee ACI 209, II (chaired by Branson DE) (1971) Prediction of creep, shrinkage and temperature effects in concrete structures. Designing for creep, shrinkage and temperature. ACI SP-27, Detroit, pp 51–93
49. Ang AH-S, Tang WH (2007) Probability concepts in engineering planning and design, basic principles. Wiley, New York
50. Bayes T (1964) An essay toward solving a problem in the doctrine of chances. Philos Trans Lond 53:376–398
51. Deming WE (1940) Facsimile of two papers by Bayes. Department of Agriculture, Washington, DC
52. Strauss A, Wendner R, Frangopol DM, Bergmeister K (2012) Influence line- model correction approach for the assessment of engineering structures using novel monitoring techniques. Smart Struct Syst 9(1):1–20
53. Strauss A, Wendner R, Bergmeister K, Reiterer M, Horvatits J (2011) Monitoring and influence lines based performance indicators. Beton- und Stahlbetonbau, 106. Heft 4, pp 231–240