The local models that we have just reviewed are all models of order two (that is, they are specified by a covariance, a variogram or a generalized covariance) and allow us to tackle only linear estimation problems. The class $\Phi$ to which we have decided to restrict our search for an optimal estimator is, by the very fact that we have adopted one of these models, the rather sparse class of authorized linear estimators of such and such an order. But practice often raises problems whose solution requires more powerful, non-linear, estimators. Let us give two examples:

i) When we are kriging, for example, a point $x$ on the basis of the available experimental points in the moving neighbourhood $V_x$, we equip our estimation with a kriging variance that we know how to calculate. But, as we have already noted, this variance is a global parameter, not a local one. It is not particularly related to the structural peculiarities of the REV in the neighbourhood $V_x$, but represents simply the average mean square of the errors one would commit by estimating each one of the points of $S$ by means of the same configuration of experimental points. If it happens that in the neighbourhood of the considered point $x$ the REV behaves more erratically or more regularly than on the average, we may expect, on physical grounds, that the error will be larger or smaller at that point respectively. But since our estimation variance has only a global meaning, it cannot account for the above effect. Within the framework of a sliding representation it is impossible, by construction, to give a local meaning to the estimation variance. But even if we cannot localize the error (that is, express it as a function of the point $x$ to be estimated) we may still hope to be able to conditionalize it (that is, to express it as a function of the values $z_i$ observed at the experimental points $x_i \in V_x$). This conditional variance could be a satisfactory substitute for the localized variance that is inaccessible within the framework of the model. Indeed, if the REV is more or less dispersed in the neighbourhood of $x$ than elsewhere, then the experimental values $z_i$ observed in $V_x$ will themselves be more or less dispersed than on average, respectively, and consequently, if the model fits, the conditional variance may be able to account for this local effect.

ii) The second example comes from a pollution problem (there would be an analogous, but more complex, formulation for the problem of selective mining).
Having measured the pollution levels $z_\tau$ at some points $x_\tau$, we are seeking the probability that the mean pollution level

$$z(v_{x_0}) = \frac{1}{v} \int_z z(x_0 + x)dx$$

in some neighbourhood $v_{x_0}$ of a given point $x_0$ is higher than some alarm threshold $z_0$. With a sliding representation (assuming that $v \subset V$), we must therefore evaluate the probability $P[Z(v) > z_0]$ conditioned on the observations $Z_i = z_i$ available in the moving neighbourhood $V$.

These two problems, as well as others analogous to them, require conditional laws for their solution. As usual, before putting to use a mathematical concept, we must ask two questions: is it possible to redefine the concept in a purely operational way (after the fact) and can we estimate it in a reasonable manner on the basis of the actual information (in praxi)? The answer to the first question will be only partially positive. The answer to the second will be categorically negative as soon as there are more than one or two conditioning points, and the conclusion will be as follows: even when we claim and think that we are using conditional expectation, that is, that we are searching for our optimal estimator in the immensely vast class of all measurable functions, what we are doing in reality is quite different, and the space $\Phi$ of functions that we are really searching is always far more restricted. Conditional expectation, as it can be defined after the fact in operational terms, is always inaccessible in praxi (even when we harbour the illusion that we have attained it) and we always end up by replacing it (consciously or unconsciously) by expressions that are much simpler, and involve some degree of gross approximation.

**After the Fact Objectivity of Conditional Laws**

Consider a sliding representation, that is, the RF defined by $Z(h) = z(x + h)$, $h \in V$, and $x$ uniformly distributed in $S$. Let us investigate, within this strictly objective framework, the form of the various conditional laws that we might be interested in. Without entering into details of mathematical formalism, it is clear that for a given $h \in V$ and a given real number $\zeta$, fixing the value $Z(h) = \zeta$ of the RF $Z(h)$ is equivalent to forcing the generic point $x$ to range over the set \( \{ x : z(x + h) = \zeta \} \) instead of the whole of $S$. This set is the translate $L_{-h}(\zeta)$ by $-h$ of the level curve (in $R^2$) or level surface (in $R^3$) $L(\zeta) = \{ x : z(x) = \zeta \}$. Therefore, after conditioning, the generic point $x$ is no longer uniformly distributed in $S$, but has a conditional distribution concentrated on the curve $L_{-h}(\zeta)$. If now $f$ is a random variable of the model, that is, a function measurable on $S$, its conditional expectation is no longer $(1/S)\int f(x)dx$. It is now given by an integration over the curve $L_{-h}(\zeta)$. 