Chapter from the book *Principles, Application and Assessment in Soil Science*

Downloaded from: http://www.intechopen.com/books/principles-application-and-assessment-in-soil-science

Interested in publishing with IntechOpen?
Contact us at book.department@intechopen.com
An Application Approach to Kalman Filter and CT Scanners for Soil Science

Marcos A. M. Laia\textsuperscript{1,2} and Paulo E. Cruvinel\textsuperscript{1,2}

\textsuperscript{1}Embrapa Instrumentation
\textsuperscript{2}Physics Institute of São Carlos, University of São Paulo
Brazil

1. Introduction

Over the years, experts on soil science have brought together researchers from various fields with the aim of pooling efforts in order to characterize the properties of soils. The use of these results in agriculture through various activities can be directed towards a development whose natural resource base can be maintained in a long term. Conservation, minimization of soil pollution together with the development of irrigation and more efficient and more cost effective drainage systems optimize the efficiency of water use and nutrients in agricultural production.

The soil can be preserved through management methods, which seek to prevent deterioration, induced erosion and deposition of sediments into rivers. Environment degradation can be caused naturally and by humans. The way in how to use the soil can cause degradation and excessive compression, salinisation and acidification.

Several conventional techniques have been used to find answers to the various physical mechanisms and also the chemical and biological processes that occur in soils. Among those techniques are included: neutron probe, gravimetry, direct transmission of rays, plotters, microscopy and mercury intrusion, however these mechanisms have limitations.

New techniques that enable an accurate prediction aim at managing the flow of contaminants through unsaturated soil zone. With the goals of acquiring non-invasive samples and reaching higher resolutions on $\gamma$ and x-rays computerized tomography (CT) and on Nuclear Magnetic Resonance (NMR), which provide cross-section images from the analysed objects. NMR, however, presents strong restrictions for its use in porous media that contain paramagnetic materials (Crestana & Nielsen, 1990) and, besides this, it is difficult or even impossible to quantify the results by correlating the NMR signal and the content of water. However, by means of CT it is possible to figure out a good correlation between the x-ray linear attenuation coefficients and the water content in soils.

The quality of an image is one of the key requirements for its analysis and it is desirable that the reconstructed object is very close to the tested sample. The use of algorithms developed for the areas of human knowledge has grown considerably and has improved image processing for visual information analysis and human interpretation or the automatic perception of machines. In human interpretation, x-ray images are used not only in medicine, but also in geology, in archaeology, and in soil science for agricultural purposes.
On the other hand, the perception of automatic machines for present automatic recognition of faces, characters, fingerprints, computer vision, control of robots for surveillance, automatic processing of satellite images for fire recognition, climate change and identification of storms and hurricanes. Image processing aims at modeling the characteristic of the human eye, study processed images, such as Fourier transform and other separable image transforms. It also allows designing filters used to retrieve an image and use masks so that the processed image is more applicable than the original (real).

This study aims at improving the quality of tomographic images by filtering the signals before their reconstruction and reaching an image quality in the reconstructed slice obtained by the projections close to the real one. Digital processing algorithms can be used to work with the image by using computer vision techniques such as segmentation (for feature extraction) and Hough transform (used in order to detect geometry, ie, detection of pores), which allow the count of pores in the soil. Classification algorithms can also be used to characterize the type of soil as well as its chemical components based on the values of pixels and density.

Both kinds of algorithms can be an auxiliary tool for a reliable analysis on the impacts of the use of agricultural machinery on the soil, which result in the increase of soil density due to the compression caused by the machinery axis. This increase is directly linked to the reduction of larger diameter pores that cause a decrease in water and in nutrients and also to the gas diffusion and root penetration.

Besides, this chapter aims at presenting the use of unscented Kalman filter (UKF) and the algorithm used to separate a noise from a signal. This will be done by showing that the unscented Kalman filter together with artificial neural networks is the best option for filtering. There will be an overview and a specification after each equation of the algorithm. Moreover, this study complemented previous works that aimed at creating a better algorithm by modeling and testing results with different types of Kalman filter and with different approaches in the physical model and kinds of noise (Laia et al., 2007; Laia & Cruvinel, 2008a; Laia et al., 2008b; Laia & Cruvinel, 2009; Laia & Cruvinel, 2011), which also present a comparison with an artificial neural network (ANN) solution.

2. Computerized Tomography scanner

In radiology, computed tomography (CT) consists of an image that is derived from computerized processing of data obtained from a series of x-ray angled projections which reproduce a cross-section (a "slice") of the object under study. CT, such as conventional radiology, is based on the fact that x-rays are partially absorbed by various materials. While materials like plastic and water are easily traversed by x-rays, others, like metals, are not. This technique had already been widely applied to medical areas, however its use in soil science was introduced by Petrovic (Petrovic et al., 1982), Hainsworth and Aylmore (Hainsworth & Aylmore 1983) and Crestana (Crestana, 1985). Petrovic made it possible to use x-ray CT to measure the density of soil volumes, while Crestana demonstrated that CT can solve problems related to studies of the water physics in the soil. From these studies, a project that developed a scanner for soil science has been created (Cruvinel, 1987; Cruvinel et al., 1990; Cruvinel et al., 2009).
2.1 Computerized Tomography in soil science

The application of CT in soil science to investigate the same physical properties plays an important role in studying the transportation of water and solutes within this environment. The direct transmission of $\gamma$ or x-rays provides a major contribution in order to solve the various problems in the field of soils with results on a scale of the order of millimeters, however many responses are still expected on the level of particles, macropores and micropores.

Figure 1 illustrates different CT scanners that are dedicated to soil science and were developed and installed at Embrapa Instrumentation (São Carlos, SP - Brazil), which are based on sources of $\gamma$ and x-rays for the study of soils and plants and also allow the use of various sources of radiation and intensities of energy. Several studies have been developed to improve the visualization of the images acquired and the reconstruction algorithm, as well as the equipment developed (Venturini, 1995; Minatel, 1997; Granato 1998; Mascarenhas et al, 1999).

Fig. 1. Scanners developed at Embrapa Instrumentation: a. mini CT scanner developed for agricultural applications (Cruvinel, 1987); b. portable CT scanner for field use (Naime, 1994); c. tomography with micrometric precision (Macedo, 1997) and d. Compton scattering scanner (Cruvinel & Balogun, 2006)

Compared to classical methods, such as the direct transmission of $\gamma$-ray and gravimetric tests, the CT scanner presents the advantage of measuring a) heterogeneities in the soil b) the density of the soil and c) the moisture content pixel by pixel, and also gets a two-dimensional picture or three-dimensional soil samples from non-invasive and independent of the geometry and shape, even by using different energies and radioactive sources (Cruvinel et al. 1990).
Basically, a CT indicates the amount of radiation absorbed by each portion of the analyzed section, translates these variations into gray scale and also produces an image. As the x-ray absorption capacity of a material is closely related to its density, different density areas will present different intensities, which can be seen in monochromatic color and if you apply a mask of pseudo-color, it allows their distinction more clearly. Thus, each signal value corresponds to the average absorption of the tissue in the area, expressed in Hounsfield units (named after the creator of CT machines). Each projection represents an average and each set is stored in a projection matrix. Based on the intensity emitted by the x-ray source and the intensity captured by the detector at the other end of the propagation line, one can determine the attenuation weight due to the object that is located between the source and the detector. The data on the attenuation weight is crucial to the reconstruction process and enables a mapping of the linear attenuation coefficient of the object cross section. This coefficient mapping is represented by pixels whose values are given by the so-called CT Numbers. These numbers are normalized according to the water attenuation coefficient $\mu_{\text{water}}$. In other words, the CT numbers are defined in equation 1:

$$CT\ Number = \frac{\mu - \mu_{\text{water}}}{\mu_{\text{water}}} \times 1000$$  \hspace{1cm} (1)

where $\mu$ is the linear attenuation coefficient of the analyzed body.

From this number, it is possible to obtain a map of the attenuation coefficients, which allows a more detailed analysis of the body being studied. In medicine, it was agreed that the water CT number is equivalent to 0 (zero).

The integration function of the object along the ray is represented by a line integral. Each set of line integrals of parallel rays form a parallel projection (Figure 2), which can be treated having the signal theory as its base.

The main advantage of CT is to allow the study of cross sections without intrusion, that is, an unparalleled improvement in relation to techniques of soil analysis, which, in general, are invasive and can destroy important features that could be preserved. It is important to obtain a perfect image quality that prevents materials found in the soil from being erroneously interpreted. This is achieved when the material in focus is the same found by an intrusive manual search.

Fig. 2. Two parallel projections of an object expressed by a two-dimensional function
2.2 Noise that can occur in CT
In CT there are three main processes related with the interaction of the radiation with the matter, which are the photoelectric effect, the Compton effect and pair production effect (Cruvinel et al. 1990). Besides the issues that are related to the effects arising from the energy range used in the source, others also influence the measurement in computed tomography, like the statistics of photon counting. The probability of detecting photons in a range of exposure time $t$ can be estimated by the distribution function of Poisson (Deremack & Crowe, 1984), (Cruvinel, 1987).

$$P_s(\lambda, t) = \frac{\lambda^n}{n!} e^{-\lambda},$$

(2)

where $\lambda$ is the number of photons and $\bar{\lambda}$ is the average number of photons emitted in the time interval $t$ and $n$ indicates the efficiency of the photomultiplier, as shown in the following expression:

$$\bar{\lambda} = \xi Mt$$

(3)

where $M$ is the average ratio of photons (photons / second) and $\xi$ is the quantum efficiency of the photomultiplier. The uncertainty or noise is given by the standard deviation

$$\sigma = \sqrt{\bar{\lambda}} = \sqrt{\xi Mt}.$$ 

(4)

Therefore, the signal to noise ratio (SNR), presented by the incident signal, is:

$$SNR = \frac{\bar{\lambda}}{\sigma} = \frac{\lambda}{\sqrt{\xi Mt}}.$$ 

(5)

In this ratio, it is estimated that, for a small number of photons, there can be considerable noise, however, as $\bar{\lambda}$ increases, the noise might be negligible. The thermionic emission of electrons in the photo multiplier cathode can cause an increase in the noise. Considering that the photo-cathode emits electrons randomly due to the cathode current, a new signal to noise ratio equation is given by equation 7:

$$SNR = \sqrt{(\xi M + M)t}$$

(6)

$$SNR = \frac{\sqrt{\xi Mt}}{\sqrt{\xi M + M}}.$$ 

(7)

In the display of a tomographic image, there is granularity, which is significant for viewing low contrast objects. During the analysis of a CT image, there is presence of granularity, which is significant for viewing low contrast objects. The term noise in tomography images refers to the variation of the attenuation coefficient at the average value and it is obtained when an image of a uniform object is used. (Hender, 1983). The image noise can be based on the calculus of the standard deviation and also on the Weiner power spectrum of the noise, which is seen as a function of spatial frequency. In other words, it allows observing the intensity and the type of noise that involves the system also influences the image obtained.

The noise in CT images involves rounding errors in the reconstruction program (algorithm noise), electronic noise and noise caused by the display system. The main source of noise in
CT images is the change in quantum (*quantus mottle*), defined as the statistical spatial and the temporal variation in the number of x-ray photons absorbed by the detector. The noise of the algorithm depends on the pixel size of the display device and also influences the image noise, since it leads to larger pixels to reduce the image noise, however with less quality resolution. Reconstruction algorithms typically make use of anti-aliasing filters to minimize the visual effect of the noise but also present some loss in the spatial resolution.

The electronic noise can be generated by non-ideal electronic devices such as non pure resistors and capacitors, non ideal contact terminals, leakage current of transistors, Joule, and it may also be independent of the signal, due to external interference (electrical or mechanical) (Ziel, 1976).

Besides the noise, the CT images are subject to various artifacts and distortions as polychromatic sources (non-ideal or monochromatic), caused by the following effects: beam hardening, aliasing, different materials in the same voxel (partial volume), displacement of the sample or of the equipment (Duerinckx & Macoski, 1978; Joseph & Spital, 1978; Ibott, 1980; Granato, 1998).

The low-pass and the median filters can be used to solve problems of signal noise, however they can cause a crucial losses of informations. Systems having different sources of noise do not present an optimal solution when using these types of filters.

### 3. The Kalman Filter

The Kalman Filter is a mathematical tool, created more than 30 years ago, which is widely used to solve statistical problems. It is considered a good estimator for a large class of problems and an effective and useful estimator for other classes. It has recently been used in computer graphics for applications involving the simulation of musical instruments in virtual reality (VR), for the reading of speaker’s lips in video sequences, among other applications (Pereira, 2000). In 1960, Rudolf Emil Kalman published a paper describing a recursive solution to the problem of linear filtering for discrete data (Kalman, 1960).

Stanley F. Schmidt, who worked on the Apollo project at NASA, has been the first to apply it in a practical way, since he aimed at taking a spacecraft to the moon and bring it back. For this, he had to solve problems in the estimation of trajectories and control. Schmidt worked with what would be the first full implementation of the Kalman filter and it had the same integral control system of Apollo and from this experience, it has been used in most onboard estimation systems and also in trajectory control of aircrafts.

The Kalman filter is considered to be an advance concerning the estimation theory. It is used, for example, to estimate a linear-quadratic Gaussian problem, which addresses the difficulty of state estimation of a dynamic linear system disturbed by white Gaussian noise, by using measurements that are linearly related to these states and also corrupted by this noise. This filter consists of a set of mathematical equations that provide a significant and efficient computational solution (recursively) to estimate the state of a process, so that the mean square error can be minimized. The filter allows that states (past, present and even future) are estimated and may do so even when the precise nature of the modeled system is unknown (Welch & Bishop, 2004).

Several models can be found, such as: increased, extended, dual estimation, joint estimation, unscented, among others.
3.1 Discrete Kalman filter

The process to be estimated solves the general problem of estimating a state $x$, which is a process controlled in a discrete time and generated by a linear stochastic differential equation, i.e:

$$x_k = Ax_{k-1} + Bu_k + n_{k-1},$$  \hspace{1cm} (8)

with a measurement $z$, which is

$$z_k = Hx_k + v_k.$$  \hspace{1cm} (9)

The random variables $n_{k-1}$ and $v_k$ represent the process noise and the measurement (respectively). They are assumed to be independent (between them) and are considered white noise with the normal probability distributions:

$$P(n) \sim N(0, Q)$$  \hspace{1cm} (10)

$$P(v) \sim N(0, R)$$  \hspace{1cm} (11)

In practice, the covariance matrixes of the process noise (covariance $Q$) and of the measurement noise (covariance $R$) can change at each measurement.

The discrete Kalman filter (DKF) estimates a process by using a kind of feedback, in which it estimates the process state at some time and then obtains feedback in the form of (noisy) measurement. Thus, the equations can be divided into two stages: the ones to update the time and the ones to update the measures.

The equations that update the time are responsible for projecting an a priori (in time) estimate of the current state and of the error covariance to obtain an a priori estimate for the next step.

The equations that update or correct measures are responsible for the feedback - for example, the incorporation of a new measure into the a priori estimate for an a posteriori estimate.

The equations that update or predict the time can also be considered prediction equations, while the equations that update the measures can be treated as correction equations. Thus, the estimation algorithm is a predictor-corrector algorithm for solving numerical problems.

Time update equations (predictor)

$$\hat{x}_k^- = A\hat{x}_{k-1} + Bu_k,$$  \hspace{1cm} (12)

$$P_k^- = AP_{k-1}A^T + Q.$$  \hspace{1cm} (13)

Note that these equations refer to time and covariance estimation over time from step $k - 1$ to step $k$.

Measurement update equations (corrector)

$$K_k = P_k^- H^T (HP_k^- H^T + R)^{-1}$$  \hspace{1cm} (14)

$$\hat{x}_k = \hat{x}_k^- + K_k(z_k - H\hat{x}_k^-)$$  \hspace{1cm} (15)

$$P_k = (I - K_k H)P_k^-.$$  \hspace{1cm} (16)

The first task during the update of the measure is to compute the Kalman gain, $K_k$. The next one is to measure the update process and find the $z_k$ value and, then, generate an a
posteriori state estimate by incorporating it as in equation 15. The final step is to obtain an estimated covariance error by using equation 16.

### 3.2 Extended Kalman filter

A solution to nonlinear systems is the extended Kalman filter (EKF) (Welch and Bishop 2004). By analyzing the prediction function of the Kalman filter, illustrated in equation 12, it is possible to observe that the filter behaves linearly. By applying a nonlinear function, one can obtain a good prediction for the next states, the extended Kalman filter. This algorithm applies the Kalman filter to nonlinear systems simply by linearizing all nonlinear models, so that the traditional equations of the filter can be applied. The nonlinear system can be rewritten as:

\[
\begin{align*}
x_k &= f(x_k, n_k, u_k), \\
z_k &= h(x_k, n_k).
\end{align*}
\]

The algorithm is also adapted to solve nonlinear problems through the use of functions, as seen below:

**Time update equations (predictor)**

\[
\begin{align*}
\hat{x}_k^- &= f(x_k^-, 0, u_k) \\
P_k^- &= F_k P_{k-1} F_k^T + Q.
\end{align*}
\]

**Measure update equations (corrector)**

\[
\begin{align*}
K_k &= P_k^- H_k^T (H_k P_k^- H_k^T + R)^{-1} \\
\hat{x}_k &= \hat{x}_k^- + K_k (z_k - h(x_k^-, 0)) \\
P_k &= (I - K_k H_k) P_k^- 
\end{align*}
\]

For the propagation of variances, one must know the Jacobian or the Hessian transition matrices and observe the state functions given by Equations 17 and 18, respectively.

### 3.3 Unscented Kalman filter

The unscented Kalman filter (UKF) is similar to the extended version (Julier & Uhlmann, 1997). The distribution of states is represented by a Gaussian random variable, however it is specified by using a minimum set of sampling points, which are chosen carefully. The sampled points capture the true mean and the covariance of a random variable and when it propagates through a truly non-linear system, it captures the mean and the covariance accurately to promote a third order estimation for any nonlinearity. Thus, this is done through the use of unscented processing.

The unscented transformation is a method used to calculate the statistics of a random variable that undergoes a nonlinear transformation (Wan & Merwe, 2000). Considering a spread of a random variable (with dimension \(L\)) through a nonlinear function \(y = g(x)\), one assumes that it must present a mean \(x\) and a covariance \(P_x\) to calculate the statistics \(y\). It should form a matrix \(X\) of \(2L + 1\) sigma vectors (with corresponding weights \(W_l\)), according to the following:
\[ X_i = \bar{x} \]  
\[ X_i = \bar{x} + (\sqrt{(L + \lambda)P_x})i, \text{ for } i = 1, \ldots, L \]  
\[ X_i = \bar{x} - (\sqrt{(L + \lambda)P_x})i, \text{ for } i = L + 1, \ldots, 2L \]  
\[ W_{o}^{(m)} = \lambda/(L + \lambda) \]  
\[ W_{o}^{(c)} = \lambda/(L + \lambda) + (1 - \alpha^2 + \beta) \]  
\[ W_i^{(m)} = W_i^{(c)} = 1/[2(L + \lambda)], \text{ for } i = 1, \ldots, 2L \]

where \( \lambda = \alpha^2(L + k) - L \) is a scalar parameter. The variable \( \alpha \) determines the sigma scattering point around the mean \( \bar{x} \) and is always a minimum positive value. \( k \) is a secondary scalar parameter defined as 0 and \( \beta \) is used to incorporate the a priori knowledge of the distribution of \( x \) (for Gaussian distributions, \( \beta = 2 \) is optimal). \( (\sqrt{(L + \lambda)P_x})i \) is the \( i \)\-th line of the square root matrix. The sigma vectors are propagated through the nonlinear function:

\[ y_i = g(X_i) \text{ for } i = 0, \ldots, 2L \]

The mean and the covariance \( y_i \) are approximated by using the mean and the covariance of the sample (Figure 3) of the posterior sigma points

\[ \bar{y} \approx \sum_{i=0}^{2L} W_i^{(m)} y_i \]  
\[ P_y \approx \sum_{i=0}^{2L} W_i^{(c)} \{y_i - \bar{y}\} \{y_i - \bar{y}\}^T. \]

Fig. 3. The unscented transform

This method differs from the general methods of sampling (the Monte-Carlo method such as particle filters), which requires orders of magnitude with more sampling points as an attempt to define and propagate the state (possibly non-Gaussian) distributions. The unscented approaches result in more accuracy for the third order of Gaussian inputs for all nonlinearities.
The unscented Kalman filter is a direct extension of the unscented transform for the equation recursive estimation

\[ \hat{x}_k = \hat{x}_{k-1} + K_k \cdot [y_k - \hat{y}_k] \]  

(33)

where the state of the random variable is redefined with the concatenation of the original states and the noise:

\[ x_k^a = [x_k^T v_k^T n_k^T]^T. \]  

(34)

The selection of sigma points is applied to a new random variable state to select and calculate the corresponding sigma matrix \( x_k^a \). The unscented Kalman filter is pointed out by equations 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, and 45.

The Jacobian or the Hessians matrices do not need to be calculated and the calculation total numbers are the same of the extended filters related to the nonlinear controls that require feedback from the states. In these applications, the dynamic model is a physically-based parametric model, which is assumed as known.

Unscented Kalman filter algorithm

Sigma points calculation:

\[ X^a = [\hat{x}_{k-1}^a \hat{x}_{k-1}^{a-1} \pm \sqrt{(L + \lambda) P_{k-1}^a}] \]  

(35)

where \( X \) is the set of points with unscented transformation based on the mean and on the \textit{a priori} covariance.

Prediction equations:

\[ \hat{x}_k^- = \sum_{i=0}^{2L} W_i^{(m)} x_i^{k|k-1} \]  

(36)

where \( W^{(m)} \) represents the set of sigma point weights used for true mean reconstruction.

\[ P_k^- = \sum_{i=0}^{2L} W_i^{(c)} [x_i^{k|k-1} - \hat{x}_k^-] [x_i^{k|k-1} - \hat{x}_k^-]^T \]  

(37)

where \( W^{(c)} \) represents the set of sigma point weights used for true covariance reconstruction.

\[ X_{k|k-1}^c = F(X_{k-1}^x, X_{k-1}^v) \]  

(38)

where \( F \) is the function of sigma propagation for state transitions.

Correction equations:

\[ Y_{k|k-1} = H(X_{k|k-1}^x, X_{k|k-1}^v) \]  

(39)

where \( H \) is the system function of sigma point generation for observation states \( Y \).

\[ \hat{y}_k = \sum_{i=0}^{2L} W_i^{(m)} Y_i^{k|k-1} \]  

(40)

where \( y \) is the observed state estimation reconstructed for the sigma points.

\[ P_{\hat{y}_k \hat{y}_k} = \sum_{i=0}^{2L} W_i^{(c)} [Y_{i,k|k-1} - \hat{y}_k] [Y_{i,k|k-1} - \hat{y}_k]^T \]  

(41)

\[ P_{x_k \hat{y}_k} = \sum_{i=0}^{2L} W_i^{(c)} [X_{i,k|k-1} - \hat{x}_k] [Y_{i,k|k-1} - \hat{y}_k]^T \]  

(42)
where $K$ is the Kalman gain obtained through the noisy covariance

$$\hat{x}_k = \hat{x}_k^- + K(y_k - \hat{y}_k^-)$$

this equation represents the correction of a priori and

$$P_k = P_k^- - K(P_{\hat{y}_k}y_k)K^T$$

represents the correction of the a priori covariance.

The Kalman filter was originally designed to solve a problem of state estimation and has been used in many applications. This newer filter also provides a better performance. Thus, the unscented Kalman filter and the extended filter present the same order of complexity.

Due to the numerical instability of the noise and the use of the Cholesky factorization to determine the square root of the probability matrix, van der Merwe and Wan have developed the Square-Root Unscented Kalman Filter (SRUKF) (Merwe & Wan, 2001), which allows a better control of the variance matrix values and bypasses the problem of becoming a negative or indefinite matrix. As the original unscented Kalman filter, the square root filter is initialized by calculating the square root of the covariance matrix states by the Cholesky factorization:

$$S_0 = \text{chol}[E[(x_0 - \hat{x}_0)(x_0 - \hat{x}_0)^T]]$$

However, the spread factor and the Cholesky update are then done in subsequent iterations to directly form the sigma points. In the equation below, the update time of the Cholesky factor is calculated by using a QR decomposition of the matrix composed of the propagated sigma point weights and the square root of the covariance matrix of the additive noise case:

$$S_k^- = qr([\sqrt{W_1^{(c)}(X_{1:2L,k}^- - \hat{x}_k^-)\sqrt{R^T}]}$$

A subsequent Cholesky update (or regression) in the equation below is needed since the weight zero is perhaps negative:

$$S_k^- = \text{cholupdate}\{S_k^-, X_{0,k}^-, \hat{x}_k^-, W_0^{(c)}\}$$

These two steps replace the time update. They are also used to calculate the Cholesky factorization and the error covariance of the observation:

$$S_{\hat{y}}_k = qr([\sqrt{W_1^{(c)}(Y_{1:2L,k} - \hat{y}_k^-)\sqrt{R^T}}]$$

$$S_{\hat{y}}_k = \text{cholupdate}\{S_{\hat{y}}_k^-, Y_{0,k} - \hat{y}_k^-, W_0^{(c)}\}.$$ (50)

Unlike the way in which the gain of Kalman filter is calculated in standard unscented, it is calculated here by using two inversions:

$$K_k(S_{\hat{y}}_k S_{\hat{y}}_k^T) = Px_k y_k$$

Since it is square and triangular, efficient replacements can be used to solve it directly, without having to invert the matrix. Finally, the Cholesky factorization updates the state
covariance in the equation below and it is calculated by applying sequential Cholesky regressions:

\[ S_k = \text{cholupdate}(S_k^-, U, -1) \]  

(52)

The vectors are represented by the columns of the equation regression 53. This update replaces the previous equation 45:

\[ U = K_r S^\_y \]  

(53)

By knowing the process function and with a Kalman filter that supports non-linear functions, it is possible to have a significant improvement in the signal. Another solution is to use a neural network to promote a better function of the mapping process by reducing the projections noise. For an estimation of the neural network weights together with the state estimates, two methods of filtering can be used: the estimation and the dual estimation. These arrangements are used to determine the filtering when the initial weights are known and the next state is obtained through a linear mapping as the previous one.

### 3.4 Joint estimation

Since the transfer function is not known and with the aim of increasing the filter order, a new estimation method has been used, in which it is possible to estimate new parameters from the states of the hidden chains of the Markov model. The main problem involves the identification of functions required to estimate states and parameters. The prediction equations can be described as:

\[ x_k = f(x_{k-1}, n_k) \]  

(54)

\[ z_k = h(x_k, v_k) \]  

(55)

The parameter estimation involves the determination of a nonlinear mapping

\[ y_k = g(x_k, W) \]  

(56)

where \( x_k \) is the input, \( W \) is the weight and \( y_k \) is the output. The nonlinear mapping \( g \) is parameterized by the vector \( W \). The nonlinear mapping can be done by an artificial neural network. The learning corresponds to estimating the parameters of \( W \). Training can be done with pairs of samples, consisting of a known input and a desired output \((x_k, d_k)\). The machine error is defined by equation 57. The learning objective is to minimize the expected squared error.

\[ e_k = d_k - g(x_k, W) \]  

(57)

The UKF can be used to estimate the parameters by using a model for network training that writes a new state-space representation:

\[ w_k = w_{k-1} + v_k \]  

(58)

\[ y_k = g(x_k, w_k) + e_k \]  

(59)

where the parameter \( w_k \) correspond to a stationary process with an identity matrix of state transition, governed by a procedural noise \( v_k \) (the choice of variances determines the
filtering performance). The output $y_k$ corresponds to a nonlinear observation $w_k$. The extended Kalman filter can be applied directly as an efficient technique for the correction of second order parameters. As the problem in focus is to work with unobserved input $x_k$ and requires an estimation of states and parameters, one should consider a dual estimation problem, taking a dynamic and discrete-time nonlinear system into account

$$x_{k+1} = f(x_k, v_k, W)$$

(60)

$$z_k = h(x_k, n_k)$$

(61)

where states are $x_k$ and the set of parameters $W$ of the dynamic system should be estimated only from the noisy signal $y_k$.

The dynamic system can be seen as a neural network, where $W$ is the set of weights and function $f$ corresponds to a function of neural network that uses an input $x_k$. Thus, by applying these equations to the unscented Kalman filter, one can have a new function for estimating and observing new states. One approach to neural networks can be seen in (Laia & Cruvinel, 2008b; Laia & Cruvinel, 2009).

4. System modeling

The physical model of photon counting is defined by the equation:

$$I_k = I_0 e^{-\mu_d}$$

(62)

where $I_0$ is the number of photons leaving the source, $\mu$ is the material degree of absorption, $d$ is the distance between the source and the detector and $I$ is the number of photons that cross the material and reach the detector. The counting of photon is affected by a Poisson-type noise. For a closer model to the physical one, each projection is analyzed individually, as if they were in time-varying positions. This classical approach enabled to develop a dynamic estimation of noise-free projections (Figure 4):

$$I_{p_k} = f(I_{p_{k-1}}, q_{k-1})$$

(63)

$$I_{s_k} = h(I_{p_k}, r_k)$$

(64)

where $I_p$ is a noise-free projection and $I_s$, a projection disturbed by noise. The variables $q$ and $r$ represent the white noise, ie, present distribution $q \sim N(0, Q)$ and $r \sim N(0, R)$, respectively. Function $f$ can be used as a mapping neural network or as a transfer state matrix. Function $h$ hides the unobserved states and can also be an array. It can also be adjusted by using the Poisson noise and the Anscombe transform. Some of the previous studies that focused on this approach and obtained good results are (Laia et al, 2007; Laia & Cruvinel, 2008b; Laia & Cruvinel, 2009).

A filtering proposal is needed for a new model (based on the physical one) to determine the process variables and how the observation is carried out. The equation of a process defines the previous state $x_{k-1}$, and through a transformation influenced by a function $f$ and white noise $q_k$, one can reach a new state $x_{k-1}$. These states can be hidden from the system output. Thus, it is possible to define a new function $g$ that transforms this variable according to what is observed.
The uncertainty estimation process of the current value has a confidence interval that consists of photon counting and is given by Poisson noise. This uncertainty can be filtered with an estimation of other measures of the equation that are independent on the confidence interval. Thus what ends up being filtered is the noise from the detector, i.e., mechanical and electronic noise.

In order to increase confidence and get more reliable values, one can change the focus on what is being estimated.

The different distributions can be mapped with specific non-linear transformations that alter the Gaussian distributions, such as the Anscombe (Poisson) or the inverse Box-Muller (Uniform), however they are linear approximations that present cumulative errors. The Kalman filter is limited to solving stochastic problems, given by the following equations:

\[
\mu_k = f(\mu_{k-1}, q_k) \tag{65}
\]

\[
I_k = h(I_0, \mu_0, d, \eta_k) \tag{66}
\]

Equation 65 is the transition function that updates the measure of the attenuation coefficient or the material degree of absorption while variable \( q \) is the uncertainty of this process equation. Equation 66 is the observation function that generates the final projection from the degree of absorption and \( d \) the distance between the source and the photon detector. Variable \( r \) is the noise that is transferred to the projection. This model can be approximated by a second order filter in order to estimate the variables of the process equation and to use the Anscombe transform to the observation equation:
\[
\begin{bmatrix}
\mu_k \\
\Delta \mu_k
\end{bmatrix} = \begin{bmatrix}
1 & 1 \\
0 & 1
\end{bmatrix} \begin{bmatrix}
\mu_{k-1} \\
\Delta \mu_{k-1}
\end{bmatrix} + \begin{bmatrix}
0 \\
q_k
\end{bmatrix}
\]

(67)

\[I_k = h(I_0 e^{-\mu_k d}, r_k)\]

(68)

Function \( h \) is related to the Anscombe transform (Anscombe, 1948), which transforms a Poisson noise into a Gaussian noise, with variance value close to 1. As the noise variation presents a Poisson distribution, the mean and the variance of values are equivalent to the counting of photons. A new system can be defined based on an equation, in which the ray sum \( \mu_k \) is the variable used in the process (Equation 76) and it also allows the observation of the projection, as it is presented in the array of projections:

\[
\mu_k = f(\mu_{k-1}, u_k, q_k)
\]

(69)

\[I_k = h(I_0 e^{-\mu_k d}, r_k)\]

(70)

Variable \( u_k \) consists of an external input that is related to the prediction of new states. In order to promote a better estimation of the states without noise, one can use neural network to determine the behavior of the equation of the process that uses the Kalman filter for a dual estimation:

\[
\mu_k = f(\mu_k, w_k, q_k)
\]

(71)

\[I_k = A(I_0, r_k)e^{-\mu_k d}\]

(72)

\[w_k = w_k + a_k\]

(73)

\[e_k = g(\mu_{k-1}, w_k) - \mu_k\]

(74)

where \( a_k \) is used to update the weight and \( e_k \) indicates the error between inputs and the desirable outputs.

To promote a better estimation of the transition states of the process, an artificial neural network can be used:

\[
\begin{bmatrix}
\mu_k \\
w_k
\end{bmatrix} = \begin{bmatrix}
g(\mu_{k-1}, u_k, W_{k-1}) + q_k \\
w_{k-1} + a_k
\end{bmatrix}
\]

(75)

\[I_k = h(I_0 e^{-\mu_k d}, r_k).\]

(76)

Now, by focusing on the observation equation, the observation noise variance can be treated with the Anscombe transform through the use combined with its inverse.

\[I_k = A^{-1}(A(I_0 e^{-\mu_k d}) + r_k)\]

(77)

where \( A \) represents the transformation and its inverse \( A^{-1} \). This makes it possible to work with a Gaussian noise with distribution \( r \sim N(0,1) \).

In this model, the input has been used as the current observation of the system (ie, the attenuation coefficient of the noise) so that one could take advantage of the neural network functionality which provides a nonlinear mapping. The neural network itself consists of the interaction between the value measured before and the noise presented now, in order to better predict the data.
An equation based on equation 5 can be developed in order to determine the Poisson noise given by the following expression:

$$ I \pm \sqrt{I} . \quad (78) $$

Thus, the observation equation can be written as follows:

$$ I_k = I_0 e^{-\mu_k} + \sqrt{I_0 e^{-\mu_d}} r_k . \quad (79) $$

where \( r \), since it presents a Gaussian distribution, can assume a negative or zero value. This alternative is allowed to deviate what is considered to be only an approximation from the use of the Anscombe transform.

Another model for the equation could be simplified as:

$$ I_k = I_0 e^{-\mu_d} + r_k . \quad (80) $$

where \( r \) is replaced by a variance given by the signal noise ratio of

$$ \frac{I_k}{\sqrt{I_k}} . \quad (81) $$

This approach allows the inclusion of other noises in the system of photon counting through the propagation of errors. In order to set the noise variance \( R \) being treated, one has to take the number of primary quantities into account. It is measured from the observation of the system variables \([I_0, \mu] \). The value of \( I \) depend on the relationship between these variables. In formal language, that is:

$$ I = G(I_0, \mu, d) . \quad (82) $$

In case the errors with the measured magnitudes \( I_0 \), \( \mu \) and \( d \) are \( \Delta I_0 \), \( \Delta \mu \) and \( \Delta d \), the photon counting error \( \Delta I \) is given by the expression:

$$ \Delta I = \left| \frac{\partial I}{\partial I_0} \right| \Delta I_0 + \left| \frac{\partial I}{\partial \mu} \right| \Delta \mu . \quad (83) $$

The values of \( \Delta I_0 \) and \( \Delta \mu \) are given by mean standard deviation or by their estimator as there are many or few magnitude measures \( I_0 \) and \( \mu \).

When the sample size is adequate, it can determine the statistical error of independent magnitudes and these can become the dependent magnitude variance for calculating the statistical error.

$$ \sigma_I^2 = \left( \frac{\partial I}{\partial I_0} \right)^2 \sigma_{I_0}^2 + \left( \frac{\partial I}{\partial \mu} \right)^2 \sigma_{\mu}^2 + \left( \frac{\partial I}{\partial d} \right)^2 \sigma_{d}^2 . \quad (84) $$

Besides the uncertainty of the system variables, the detector presents a characteristic noise. This noise variation is known when the detector closes, in other words, as no photons are released to it, it still gives the score. As it is associated directly to an additive noise, you can add the noise variation into the equation as an error for variable \( I \):

$$ \sigma_I^2 = \left( \frac{\partial I}{\partial I_0} \right)^2 \sigma_{I_0}^2 + \left( \frac{\partial I}{\partial \mu} \right)^2 \sigma_{\mu}^2 + \left( \frac{\partial I}{\partial d} \right)^2 \sigma_{d}^2 + \left( \frac{\partial I}{\partial \sigma} \right)^2 \sigma_{\sigma}^2 . \quad (85) $$

\( \sigma_I^2 \) is used to define the noise variance \( R \) of the observation equation of the filter. Then, it is necessary to define the variance of the process. In the literature, Haykins (Haykins, 2001) has defined several ways to infer on the process variance \( Q \). Since the signal is observed, the
noisy signal variance can be used. As the process variance is directly linked at the $\mu$ vector, the equation can be changed for a noise $\mu$:

$$\mu = \ln\left(\frac{I}{d}\right) / d$$  \hspace{1cm} (86)

From this vector, a variance $Q$ needs to enter the filter. It can be obtained as follows:

$$Q = \sigma^2_\mu$$  \hspace{1cm} (87)

Another important step is the definition of the control constants: $\alpha$, $\beta$ and $\kappa$. As the process variable $\mu$ magnitude differs from the observation of variable $I$, $\alpha = 1$ has been chosen. In case this value is greater or lower than ideal, there is no possibility of filtering or causing numerical instabilities.

As suggested in the literature, for a parameter or a joint estimation filters, the variable remained $\beta = 2$, while the $\kappa$ value was given $3$ - the number of neurons.

With the aims of comparing the efficiency of the filter and setting up a filter to ensure a desirable image quality, both of the algorithms have been applied to various soil samples. The first sample consists of sand grains in a Plexiglas envelope. The second and third samples are portion of natural soil. The fourth and the fifth are present in degraded soil bulks and the sixth shows a portion of naturally cemented soil. In an artificial neural network there are two neurons in the input layer, two in the intermediate and one in the output layer.

For the inputs of the filter, the same uncertainty was used, as all of the samples have been generated in the same CT scanner. The variance $\sigma^2_\nu$ was obtained by the maximum of vector projections. After the filtering process, it has been applied to the maximum of the projection matrix. The variance $\sigma^2_\lambda$ was the same used in the process variance $Q$. For the variance $\sigma^2_\phi$, the value used was 0.05, corresponding to the uncertainty of measurement in millimeters.

For the errors in the photon counting detector, the value used was 100 for the variance $\sigma^2_\epsilon$. The results obtained from this new modeling system are presented below. In Figures 5, 7, 9, 11 and 13, it is possible to visualize the comparison between the signals of a set of projections by using the samples filtered with the SRUKF: Original projections (red), Linear estimation (blue) and non linear estimation with ANN (green). In Figures 6, 8, 10, 12, 14 and 16, it is possible to visualize the reconstructed images by using the filtered back-projection algorithm.

![Fig. 5. Comparison between the projections of sand grains](image-url)
Fig. 6. Comparison between the reconstructed images of sand grains. a. Original projections, b. Projections filtered with the SRUKF and the ANN and c. Projections filtered with linear estimation

Fig. 7. Comparison between the projections of soil bulk samples

Fig. 8. Comparison between the agricultural soil reconstructed images for the visualization of bulk information. a. Original projections, b. Projections filtered with the SRUKF and the ANN and c. Projections filtered with linear estimation
Fig. 9. Comparison between projections of the soil bulk sample.

Fig. 10. Comparison between the agricultural soil reconstructed images for the visualization of bulk information. a. Original projections, b. Projected filtered projections with the SRUKF and the ANN and c. Projections filtered with linear estimation.

Fig. 11. Comparison between the projections of degraded soil samples.
Fig. 12. Comparison between the reconstructed images of degraded soil sample. a. Original projections, b. Projections filtered with the SRUKF and the ANN and c. Projections filtered with linear estimation.

Fig. 13. Comparison between the degraded projections of a soil sample.

Fig. 14. Comparison between the reconstructed images: a. Original projections, b. Projections filtered with the SRUKF and the ANN and c. Projections filtered with linear estimation.
In the signal comparisons it is possible to see the estimation errors. These errors in linear estimation promote false artifacts and excessive anti-aliasing. Also, the excessive filtering process with linear estimation has promoted losses in the images details. The SRKUF with the ANN eliminates any noise in a dynamic way and, thus, there is a higher predominance of low values of photon counting. Every single detail is preserved in the signal filtered with the SRUKF.

The signal filtered with the ANN algorithm presents the details better, a precise correction and the interest object (the porosity of soil bulk) is preserved while the signal filtered with the linear estimation promotes excessive anti-aliasing and detail loss. The false porosity (due to granularity of Poisson noise) has been eliminated in the images as it can be observed in the outer parts of the sample. In the image obtained by the filtering with the ANN, the contrast has been preserved as well as small-pointed (critical) elements in the image.

The low-contrast in this image is changed due to the false estimation in the projections like a high value estimated by filter. The single elements are either preserved or changed a little due to the contrast. With the linear estimation, the excessive filtering provoked a decrease in the contrast, which lead to the elimination of pores and high-contrastive elements. The outlines of the objects in the image are also lost when there is an excessive smoothing.
Another important factor is the separation between porosity and granularity of the images. The granularity can generate false micro-pores or fake elements. On the other hand, excessive smoothing can hide the pores and the major elements in the soil composition.

5. Conclusions

The Kalman filtering that uses linear estimation promotes a smoothing of values and do respect the nature of data distribution (linear attenuation coefficients), which present a uniform distribution, while the filter is limited to working with a Gaussian process. By analyzing the results presented in Figures 4, 5, 6, 7, 8 and 9, it was possible to observe smoothed projections and estimation errors, while Figures 10, 11, 12, 13, 14 and 15 presented losses of important details, such as micro pores and other important elements of the soil, which enables to characterize it.

Nowadays, with the use of artificial neural networks, the results already show a better transition among the variables for the estimation process. Besides mapping the behavior of the sample data, the nonlinear function also makes the necessary transformation of the process uncertainties, which helps to preserve a greater number of details on an original image.

Besides, the regular presence of artifacts or distortions on the image were due to three factors: the limitation of the reconstruction algorithm because of the ramp filter, which is necessary to observe the contrast of different materials and soil porosity; the inaccurate choice of the noise variances because of the different resolutions of the samples obtained and due the equipment noise during data collection.

A better measurement of the noise variations in the detector should be made in real time. The closer the variance value is to the real value, the less it is prone to reconstruction errors or smoothing.

Tests conducted earlier have proved that the increase of neurons, layers or number of states has not affected the filter efficiency but the processing time was longer.

More recent researches involve the implementation of the filters shown in this study in an embedded system in order to ensure better results when filter variables are fed directly according to the state of the CT scanners.

6. Acknowledgment

This work was supported in part by the National Council for Research and Development (CNPq) under Grant 306988/2007-0 and Brazilian Enterprise for Agricultural Research (Embrapa) under Grant 03.10.05.011.00.01. Also we acknowledge the Coordination for the Improvement of Higher Education (CAPES, process 573963/2008-9 and 08/57870-9).

7. References


Our dependence on soil, and our curiosity about it, is leading to the investigation of changes within soil processes. Furthermore, the diversity and dynamics of soil are enabling new discoveries and insights, which help us to understand the variations in soil processes. Consequently, this permits us to take the necessary measures for soil protection, thus promoting soil health. This book aims to provide an up-to-date account of the current state of knowledge in recent practices and assessments in soil science. Moreover, it presents a comprehensive evaluation of the effect of residue/waste application on soil properties and, further, on the mechanism of plant adaptation and plant growth. Interesting examples of simulation using various models dealing with carbon sequestration, ecosystem respiration, and soil landscape, etc. are demonstrated. The book also includes chapters on the analysis of areal data and geostatistics using different assessment methods. More recent developments in analytical techniques used to obtain answers to the various physical mechanisms, chemical, and biological processes in soil are also present.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following:


InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821