Optimal Exposure Time in Gamma-Ray Attenuation Experiments for Monitoring Time-Dependent Densities

Ana Gonzalez-Nicolas1 · Deborah Bilgic1 · Ilja Kröker1 · Assem Mayar1 · Luca Trevisan2 · Holger Steeb3 · Silke Wieprecht1 · Wolfgang Nowak1

Received: 14 December 2021 / Accepted: 30 March 2022 / Published online: 28 April 2022
© The Author(s) 2022

Abstract
Several environmental phenomena require monitoring time-dependent densities in porous media, e.g., clogging of river sediments, mineral dissolution/precipitation, or variably-saturated multiphase flow. Gamma-ray attenuation (GRA) can monitor time-dependent densities without being destructive or invasive under laboratory conditions. GRA sends gamma rays through a material, where they are attenuated by photoelectric absorption and then recorded by a photon detector. The attenuated intensity of the emerging beam relates to the density of the traversed material via Beer–Lambert’s law. An important parameter for designing time-variable GRA is the exposure time, the time the detector takes to gather and count photons before converting the recorded intensity to a density. Large exposure times capture the time evolution poorly (temporal raster error, inaccurate temporal discretization), while small exposure times yield imprecise intensity values (noise-related error, i.e. small signal-to-noise ratio). Together, these two make up the total error of observing time-dependent densities by GRA. Our goal is to provide an optimization framework for time-dependent GRA experiments with respect to exposure time and other key parameters, thus facilitating neater experimental data for improved process understanding. Experimentalists set, or iterate over, several experimental input parameters (e.g., Beer–Lambert parameters) and expectations on the yet unknown dynamics (e.g., mean and amplitude of density and characteristic time of density changes). We model the yet unknown dynamics as a random Gaussian Process to derive expressions for expected errors prior to the experiment as a function of key experimental parameters. Based on this, we provide an optimization framework that allows finding the optimal (minimal-total-error) setup and demonstrate its application on synthetic experiments.

Article Highlights
• We study the influence of anticipated density changes and experimental setup on optimal designs for GRA measurements
• We present a methodology that finds the optimal setup (minimum error) as a function of the exposure time and other parameters

* Ana Gonzalez-Nicolas
anagna@gmail.com

Extended author information available on the last page of the article
• We provide experimentalists with a quantitative understanding of unavoidable inaccuracies and how to minimize them

**Keywords** Gamma-ray · Time-dependent density · Optimal experimental design · Raster error · Noise error · Beer–Lambert’s law · Gaussian process regression

1 Introduction

Several natural and technical processes occurring in porous media lead to time-dependent densities. With time-dependent densities, we refer to the fact that the total mass within a fixed volume (i.e. a bulk density) may change over time. Examples of such processes are the infiltration of fine-grained river sediments into a coarser sediment matrix (colmation) (Schälchli et al., 2002) and mineral precipitation or dissolution within porous rocks (e.g., Hommel et al., 2015). The induced changes in density also alter other properties of porous media. Depending on the application, these alterations are desired or undesired; in both cases, it is relevant to better understand the processes, and hence to have methods for observing time-dependent densities. In other cases, it is not the solid phase of porous media that changes the density, but the varying presence of fluids. Examples include (1) the variably-saturated flow of water through porous media, e.g., in the vadose zone in the area of soil science (e.g., van Genuchten, 1980) or (2) the multiphase flow of non-aqueous phase liquids (NAPLs) through otherwise water-saturated porous media in the area of contaminant hydrology (e.g., Pankow & Cherry, 1996).

Many existing methods for monitoring time-dependent densities are destructive, i.e., they require taking material samples. However, sampling disrupts the original conditions and processes to be monitored, hence affecting the dynamics that one would like to observe. Placing sensors into the system to be monitored may be less destructive, but still, it is invasive and affects the processes at work—again disturbing the dynamic one would like to monitor. Therefore, approaches based on the attenuation of radiation energy (such as gamma rays of x-rays) are very promising. These methods have proven very useful in a wide range of environmental applications (Oostrom et al., 2007; Werth et al., 2010).

In this study, we focus on monitoring time-dependent densities via gamma-ray attenuation (GRA). We restrict this to GRA fixed in location and angle. This form of GRA experiments does not attempt to resolve spatial variations in density, but merely returns integral information on bulk attenuation (and hence on bulk density) along the beam line. Hence, our study focus should not be confused with tomographic approaches based on gamma rays that yield images, but we will discuss the extension to images and time-variable images in our outlook.

Several studies have used GRA to measure time-dependent densities. There are GRA-based studies that monitored fluid saturations (Barth et al., 2003; Eckberg & Sunada, 1984; Ferguson & Gardner, 1962; Illangasekare et al., 1995; Nofziger, 1978). Likewise, there are GRA-based studies to determine physical characteristics of soils such as soil bulk density and porosity (Baytaş & Akbal, 2002), water content, NAPLs’ saturation, hydraulic conductivity (Moreira et al., 2011), and reservoir sediments’ bulk density profiles (Beckers et al., 2018). Two very recent studies, Mayar et al. (2020) and Mayar et al. (2022) used GRA-based monitoring for colmation processes. Other fields where GRA has been successfully
applied can be found in Gharbi et al. (2004) for a colloid deposition in porous media and Sail et al. (2011) for suffusion tests on cohesionless granular matter.

In GRA, gamma rays are sent through the system, where they are attenuated and then recorded by a detector, such as a Geiger counter or scintillator. Beer–Lambert’s law links the attenuated intensity of the emergent gamma rays (per wavelength) to the (bulk) density of the traversed materials. Details for handling bulk densities of multiphasic and/or porous materials will be given in Sect. 2.1. The time allowed to count discrete gamma photons at the detector before converting them to an intensity (and then to a density value) will be called exposure time hereafter.

Both emission and attenuation of gamma rays are random processes commonly modeled as Poisson processes (Berger & Hubbell, 1987; Waggoner, 1951). Accordingly, the count numbers recorded in repeated experiments turn out to follow the Poisson distribution (Buczyk, 2009). Therefore, one can measure precise intensities only with long exposure times. In other words: the imprecision of GRA, in a static measurement scenario with constant densities, solely depends on the statistical error associated with the number of counts recorded by the detector (Mayar et al., 2020). The expected magnitude of this number, in turn, depends on the exposure time, the radiation intensity of the radiation source, and on specifics of technical equipment such as the collimators used to bundle the gamma rays through the experimental system (Mayar et al., 2020).

However, when the density to be measured changes over time and the goal is to monitor the temporal variations of density as a time series, the need for long exposure times becomes a problem: thinking in a sequence of intensities per exposure time window will lead to a temporally rastered version of the time-dependent density one wants to monitor. The temporal resolution of this rastering is directly given by the exposure time. Thus, for a given set of experimental equipment, one has to choose smaller exposure times to capture the temporal variations of density more accurately, and one enters a trade-off between precision and accuracy. Of course, one can milden the trade-off by using more powerful experimental equipment such as stronger sources or larger collimators (Mayar et al. 2019). Yet, this comes at larger costs, larger risks related to the handling of gamma sources, or reduced spatial resolution.

Our study focuses on the trade-off between accuracy and precision for given experimental equipment. We will call the underlying two types of error “noise error” and “raster error”. Noise error is caused by the statistical imprecision of gamma-ray counts. It is related to the noise part in the signal-to-noise ratio, and it decreases with larger exposure times. Raster error relates to the inaccuracy of representing a time-dependent density via constant values per exposure time. It relates to the temporal rastering quality of the signal in the signal-to-noise ratio and it increases with exposure time. Naturally, one can expect an optimum (minimum total error) for problem-specific medium values of exposure time. This raises the question: how can experimentalists find this problem-specific optimum without manual trial and error?

Our goal is to achieve minimal total errors in GRA-based monitoring of time-dependent densities by mathematical optimization of exposure time, along with other key parameters of GRA experiments. As the optimization occurs prior to the experiment, one has to work with (statistical) expectations of experimental properties and outcomes. To keep the optimization accessible to experimentalists, it should rely on prior expectations that are relatively easy to specify. Therefore, our proposed approach will only ask to specify physical parameters of the GRA device and experimental setup (Beer–Lambert parameters: incident intensity, mass attenuation coefficient, system thickness) and statistical expectations of the yet unknown time-dependent density. These statistical expectations are the anticipated
mean density, the anticipated strength of density variations, and the characteristic time scale on which the density variations are expected to occur. These specifications are sufficient to predict the magnitude of both error types as a function of exposure time and of other parameters, and hence to find the optimal exposure time (and other parameters) that minimizes the overall error.

For implementing the optimization, we will follow two different methods: a numerical computation of errors (slow and computationally expensive) and a first-order approximation (fast and still accurate). Then, we demonstrate the optimization of a hypothetical GRA monitoring exercise of time-dependent densities. We discuss the behavior of errors as a function of exposure time. Other experimental parameters may also be influenced or chosen by experimentalists, such as the strength of the gamma-ray source or the thickness of the analyzed sample. These are much less free to choose than the exposure duration, yet have impacts on experimental accuracy and precision. Therefore, we also analyze how the optimal exposure time and the achievable minimum error changes under variations in experimental conditions. The latter include both Beer–Lambert parameters (incident intensity, attenuation coefficient, system thickness, as they influence the number of counts to be expected, and hence the noise error) and the expectations on the dynamics (mean, amplitude, characteristic time, as they influence the expected signal to be rastered, and hence affect the raster error).

Finally, we also discuss how to use our framework for an overall optimization of GRA experiments for monitoring time-variable densities. The overall benefits of our work are (1) to predict controllable experimental errors in time-variable GRA prior to experimentation; (2) to provide an optimization to achieve minimal errors by choosing appropriate exposure times and (3) to assist experimentalists in the overall optimization of their experiments toward a maximum utility in experimentation and process understanding.

This paper is organized as follows. First, we explain the methodology including the system and the application of the Poisson distribution for this type of problem, the analytical and numerical approach, as well as the optimization problem. Afterward, in Sect. 3, we present our results. We also discuss overall optimization across all parameters that an experimentalist could choose freely, and discuss the limitations of our approach. Last, we conclude with a summary of recommendations for experiments and with an outlook for future research.

## 2 Methods

### 2.1 System Definition and Beer–Lambert’s Law

We consider a system and experimental setup as shown in Fig. 1. The system domain $\Omega$ has a time-dependent density $\rho(t)$ (kg/m$^3$) and a thickness $d$ (m). In the scope of our study, we neglect possible spatial variations of density. This means we assume that the experimentalist is either interested in an effective bulk density or that density can be seen as spatially homogeneous. Hence, we do not consider the density as a function of space, but only of time. For now, we provide all formulations just for a single (bulk) density to keep notation short, but we provide details on wrapping multi-phasic and/or porous media into this framework at the end of the current section.

For monitoring the density over time via GRA, the experimental system is equipped with a gamma-ray source that sends a constant incident intensity $I_0$ (before passing through
the system) through the system and measures the attenuated emergent intensity $I_1$ (after passing through the system), $I_1 < I_0$ at a photon counter (detector). The attenuation will follow Beer–Lambert’s law.

The Beer–Lambert’s law (Lykos, 1992) is an exponential absorption law that relates the attenuation of light (or other electromagnetic rays/radiation) through a substance to the properties of the substance. Beer–Lambert’s law is defined as

$$A_\gamma = \log \left( \frac{I_0}{I_1} \right) = \varepsilon_\gamma \rho d$$

and defines attenuation as the relationship between the intensities $I_0$ and $I_1$. Here, $A_\gamma$ is the system absorbance (unitless) for a given wavelength $\gamma$, $I_0$ is the incident intensity (W/m²), $I_1$ is the output intensity (W/m²), $\varepsilon_\gamma$ is the absorption concentration of the (homogeneous) material in the system (m²/kg) for the wavelength $\gamma$, $\rho$ is the concentration (here: density) of the substance (kg/m³), and $d$ is the characteristic path length (here: the thickness of the system) (m). In the following text, we will avoid all subscripts related to the wavelength $\gamma$, and assume a monochromatic source. For polychromatic sources, one would have to apply Beer–Lambert’s law to each involved wavelength. However, applying the law to all wavelengths in a lumped fashion can still be a good approximation, if $\varepsilon_\gamma$ does not vary too much with $\gamma$.

In Sect. 2.2, we will treat the intensities not as continuous variables (in units of W/m²), but as the discrete number of photon counts in a given time interval (unitless). As the fraction $I_0/I_1$ is unitless, Eq. (1) is still valid when using count numbers for $I_0$ and $I_1$.

Now, we turn toward multi-phasic and porous media. These are a volumetric mixture of several materials, each of them having its own absorption concentration $\varepsilon$ density $\rho$ and volume fraction $n$ (compare to porosity). Without loss of generality, we can look at bulk densities per constituent that already captures the volume fraction, so that we can use the thickness $d$ as identical overall sample thickness across all constituents. This is allowable due to the multiplicative character of $d$ and $\rho$ in Eq. (1). In the logarithmic form of Eq. (1), several constituents would simply add up to $(\varepsilon_1 \rho_1 d_1 + \varepsilon_2 \rho_2 d_2 + \ldots)$ in the right-hand side. Yet, GRA would only provide a single value for bulk attenuation, i.e., with the overall thickness $d$ and the individual values for $\varepsilon$ given, one would have to solve for several density values from a single piece of information. Yet, in specific cases, it is possible to disentangle them:
1. In a multiphasic or porous system, the initial state provides a base-line attenuated intensity $I_{\text{base}}$, which can be substituted for $I_0$.

2. As all volume fractions must sum up to a constant value over time, one receives an additional constraint and can solve Eq. (1) for up to two changing densities.

3. If one or more of the constituents have a negligible product of $\varepsilon \rho$ compared to all other constituents, they can be neglected in the bulk attenuation, allowing to solve for the key constituents of interest.

Such techniques are standard for GRA experiments with porous media, see e.g., Mayar et al. (2020).

2.2 The Poisson Distribution Over Different Time Windows for Counting

In GRA experiments, emerging beam intensities are typically measured by photon counts, i.e., as discrete numbers of counts per given time window. Due to the stochastic nature of gamma decay for natural sources and of attenuation by the material (Berger & Hubbell, 1987; Waggoner, 1951), the actual count is random but has a uniquely defined expected value for constant and given parameters within Eq. (1). This randomness is widely modeled via the Poisson distribution (Buczyk, 2009) because both the emitting decay process and the attenuation per emitted particle are assumed to be a collection of statistically independent events. That means the count events at the detector follow a Poisson process, such that the recorded count numbers follow the Poisson distribution (Haight, 1967).

The Poisson distribution expresses the probability of a given number of events (here: counts $X$ equaling any given value $k$) occurring in a fixed interval of time, if these events occur with a known constant mean rate and independently of the time since the last event (Haight, 1967):

$$P(X = k) = \frac{\lambda^k}{k!} \exp(-\lambda),$$

where $X$ denotes the underlying random variable, $\lambda$ is the only parameter of the Poisson distribution, and $k$ is the number of times an event could actually occur within the given time interval. Due to properties of the Poisson distribution (Haight, 1967), the parameter $\lambda$ is both the expected value $\mu_X = E[X]$ and its variance $\sigma_X^2 = \text{Var}[X]$:

$$\mu_X = \lambda$$

$$\sigma_X^2 = \lambda \rightarrow \sigma_X = \sqrt{\lambda}.$$  

In this study, $k$ is the actual number of photon counts recorded by the detector in the specified time window. Let the time window have a unit length of $L_1$ (e.g., one second to be consistent with the units used in Beer–Lambert’s law). We assume that, if all other parameters remain constant, the observed counts $k$ (recorded photon counts in the unit-length time window $L_1$) are a random outcome of a Poisson-distributed random variable $X$. Given the physics of the system, Beer–Lambert’s law can provide the corresponding expected value (Haight, 1967):
where the factor $L_1$ merely appears for unit conversion from Beer–Lambert intensities (per time) to Poisson counts (unitless number in a given time window).

A direct conversion of counts $k$ to a density value via Beer–Lambert’s law means, statistically, to use the point estimate $\lambda = E[X] \approx k$ in Eq. (4), so that one can insert $I_1 \approx k/L_1$ and then re-arrange for $\rho$. This estimate is accurate at sufficiently large values of $\lambda$, where the randomness of gamma counts can be ignored. For finite $\lambda$, however, the relative accuracy of $\lambda \approx k$, expressed as coefficient of variation $CV_X$, is:

$$CV_X = \frac{\sigma_X[X]}{E_X[X]} = \frac{\sqrt{\lambda}}{\lambda} \approx \frac{\sqrt{k}}{k} = \frac{1}{\sqrt{k}}$$

The imprecision expressed by Eq. (5) tends to zero as $\lambda \to \infty$. For smaller expected count numbers $\lambda$, the random fluctuation of the counts $k$ does matter. This is the reason why we distinguish between an actually observed count $k$, the underlying random variable $X$ that follows the Poisson distribution, and its expected value $E[X]/L_1 = I_1$, which is computable from Beer–Lambert’s law according to Eq. (4).

When choosing a time window $L$ that differs from $L_1$, one has to remember that the Poisson distribution is defined for a fixed time interval, and expresses a plain count (not a rate) without units of “per time”. Longer time windows will result, at the same count rate, in larger count numbers, but this must not be mistaken for a larger intensity or lower density according to Beer–Lambert’s law. Therefore, one has to do the conversion from counts $k$ to rates $\kappa = k/L$ (counts per time) as a data pre-processing step. Suppose we count the value $k_L$ over a time window $L(s)$ and subsequently we want to compute the rate $\kappa(1/s)$ for using $I_1 \approx \kappa$ with Beer–Lambert’s law. Then one would use

$$\kappa = \frac{k_L}{L}.$$  

The random variable underlying $\kappa(1/s)$ is denoted as $\chi$, and the random variable underlying $k_L(\cdot)$ is $X_L$. For the expected values of $\chi$ and $X_L$, this means

$$E[X_L] = \frac{L}{L_1} \cdot E[X] \approx k_L$$

$$E[\chi] = \frac{E[X]}{L_1} \approx \kappa = \frac{k_L}{L},$$

because the longer time $L$ leads to larger count numbers $k_L$, and we compensate for that fact with Eq. (6). The relevant difference arises in the imprecision. Due to the Poisson distribution, the random count $X_L$ has a variance $\sigma^2_{X_L} = E[X_L] = L/L_1 \cdot E[X]$ that is $L/L_1$ times larger than that of $X$. However, the random rate $\chi$ is not Poisson distributed, as it is merely computed from a Poisson-distributed variable $\chi = X_L/L$ (in analogy to $\kappa = k_L/L$). Instead, the division by $L$ in Eq. (6) makes its variance $L^2$ times smaller than that of $X_L$ via the plain rules of linearized error propagation. In combination, the variance of $\chi$ is $L$ times smaller than that of $X$, and can be expressed in various forms:
which confirms that longer time windows $L$ allow more precise statistical estimates of the Beer–Lambert intensity $I_1$, which will lead to more precise density estimates.

Based on the above equations, one can already construct a simple hypothesis test to specify how long of an exposure time is required to distinguish two slightly different density values by GRA in spite of the imprecision of count rates. As this test is not yet in the context of time-variable densities, it is a slight side-track to our manuscript, and therefore it can be found in Appendix D.

2.3 Errors in Monitoring Time-Dependent Densities: Raster Error and Noise Error

From now on, we consider that the density $\rho(t)$ is time-variable and, consequently in Eq. (1), we have $I_1(t)$ and $A_f(t)$. Through Eq. (4), this makes the counting process in the detector a variable-rate Poisson process with rate $\lambda(t)$. Fig. 2 shows a possible temporal variation of the rate parameter $\lambda(t)$. The red line is the theoretical $\lambda(t)$, whereas the blue crosses are observed count values $k_L(t)$ counted in discrete time windows of some length $L$. For the sake of clear illustration, we have deliberately chosen a very small $L$ compared to the total time length of the plot. This results in many blue crosses (photon counts) with high imprecision according to Eq. (5).

As already mentioned, an important parameter for designing dynamic GRA is the time window size $L$. This is the exposure time that controls the expected number of photon counts received by the detector. At the same time, the length $L$ also controls the temporal resolution of the density to be monitored, as detailed in the following.

Let us first consider the temporal resolution aspect. Basically, we cannot resolve rate variations of recorded counts within the time windows of length $L$. Instead, we will get a discretized, piecewise constant raster curve $\hat{I}_1(t)$ to approximate the theoretical intensities $I_1(t)$ at a sampling rate of $1/L$. Hence, we also would obtain a piecewise constant raster curve $\hat{\rho}_L(t)$ to approximate the targeted density $\rho(t)$. The difference between $\hat{\rho}_L(t)$ and $\rho(t)$ is an error of temporal discretization. We call the resulting error the raster error.

\[
\sigma^2 = \frac{\sigma^2}{L^2} = \frac{\sigma^2}{LL_1} = \frac{\lambda}{LL_1} = \frac{I_1}{L} \approx \frac{\kappa}{L},
\]

(9)

Fig. 2 Illustration of variable $I_1(t)$ caused by variable $\rho(t)$ in GRA monitoring. Red line shows the theoretical rate $\lambda(t)$ of the underlying variable-rate Poisson process and the blue crosses show photon counts $k_L(t)$ counted in discrete time windows of length $L$. 

© Springer
infinitely strong source, where statistical imprecision of counts would not matter, the raster error would persist for any exposure time other than zero.

In realistic setups with finite source strengths, each piecewise constant value of the raster curve \( \hat{p}_L(t) \) will also be subject to the statistical imprecision of using counts \( k_L(t) \) or count rates \( \kappa(t) \) as point estimates instead of theoretical intensities \( I_1(t) \). As this is a statistical imprecision caused by the Poisson distribution, we will call it the noise error.

The combination of raster error and noise error makes up the total error of monitoring time-dependent densities by GRA. When expressing all these errors as variances (expected squared errors) and modeling these errors to be uncorrelated, the total error \( \sigma^2_{\text{total}} \) of monitoring time-dependent densities \( \rho(t) \) is the sum of the raster error \( \sigma^2_{\text{raster}} \) and the noise error \( \sigma^2_{\text{noise}} \):

\[
\sigma^2_{\text{total}}(L) = \sigma^2_{\text{raster}}(L) + \sigma^2_{\text{noise}}(L). \tag{10}
\]

These errors will be explained in the following sections. The assumption of statistical independence between raster error and noise error is justified, as raster error is, in fact, deterministic and caused by a discretization choice of the experimentalist, while the noise error stems from the stochastic Poisson process of emission and attenuation. Of course, both errors are linked as the choice of \( L \) will jointly affect both errors, but the two errors are not directly dependent on each other, so they are statistically independent for any given choice of \( L \). In statistics, this is called conditional independence, and this is sufficient to write both errors, as in the above equation, as independent per \( L \).

### 2.4 Formal Optimization of Expected Errors

It is understood that the total error \( \sigma^2_{\text{total}} \) according to Eq. (10) is a function of the time-window size \( L \). On the one hand, large windows can reduce the noise error \( \sigma^2_{\text{noise}} \), but cannot resolve dynamics. Hence they will lead to large raster errors \( \sigma^2_{\text{raster}} \). On the other hand, small windows could resolve the dynamics well with small raster errors but suffer from larger noise errors. The optimal (minimum) total error will be achieved through an optimal window size \( L = L_{\text{opt}} \) that strikes the optimal compromise between the raster and noise error (see Fig. 3). It can be found by optimization as follows:

\[
L_{\text{opt}} = \arg \min_{L \in \Omega} \sigma^2_{\text{total}}(L), \tag{11}
\]
where $L_{\text{opt}}$ is the optimal window size. It is the argument (arg) that minimizes (min) the total error $\sigma^2_{\text{total}}(L)$ from all possible window sizes $L \in \Omega$, and $\Omega$ is the range of feasible exposure times $L$. In Sect. 1.2, we also discuss how to extend our approach toward an overall optimization across a much larger set of experimental settings.

Next, we need to formulate how exactly $\sigma^2_{\text{noise}}$ and $\sigma^2_{\text{raster}}$ depend on $L$. Before performing an experiment, the best we can do is to optimize expectations of these errors, which is exactly why we express them as variances (i.e., expected squared errors) in Eq. (10).

The Poisson distribution for varying time windows upon conversion to count rates (Eqs. (2) to (9)) gives us expectations for the noise error $\sigma^2_{\text{noise}}$, if we can anticipate (at least statistically) possible photon counts $k$ and corresponding count rates $\kappa$. This, in turn, requires anticipating possible densities $\rho(t)$ for Eq. (4).

For the raster error $\sigma^2_{\text{raster}}$, we also need some statistical anticipation of the time-dependent density. Then, we can mathematically derive expected squared errors for approximating an allegedly true density curve $\rho(t)$ by a piecewise constant, noise-affected raster curve $\tilde{\rho}(t)$. For this reason, in the upcoming section, we will treat the time-dependent density $\rho(t)$ as a Gaussian process.

### 2.5 Gaussian Processes for Time-Dependent Densities

Gaussian processes (GP) are widely used in machine learning and its applications. In the research field of geosciences, GPs are also used for interpolation methods known as Kriging (Matheron, 2019). In time series analysis, they occur as various types of autoregressive processes. From the mathematical point of view, GPs can be seen as a generalization of the well-known multivariate Gaussian probability distribution, which is completely described by its mean and (co)variance. For a comprehensive introduction to GPs, we refer to the book of Rasmussen (2006).

In our case, a GP over time, $f(t)$, can be seen as a collection of random variables indexed by time $t$, such that each finite selection of points in time results in a multivariate Gaussian distribution. A GP is completely specified by its mean function $m(t) = E[f(t)]$ and covariance function $\text{Cov}(t, t') = E[(f(t) - m(t))(f(t') - m(t'))]$. The latter describes the covariance between pairs of random variables $f(t)$ and $f(t')$ at two different points in time $t$ and $t'$. Altogether, a GP $f(t)$ is usually denoted by $f(t) \sim \text{GP}(m(t), \text{Cov}(t, t'))$.

There are many choices of covariance functions available. Most of them are parametric shapes that can be adapted to current needs via so-called hyperparameters. The most common hyperparameters are the variance $\sigma^2$, which describes the magnitude of fluctuations, and the correlation length $l$, which can be seen as the timespan that should be passed between $t'$ and $t$, before the value of $f(t)$ can differ significantly from $f(t')$. For further reading, we refer again to the book of Rasmussen (2006).

As a modeling assumption that admits statistical expectations as argued in Sect. 0, we assume that the time-dependent density $\rho(t)$ follows a GP

$$\rho(t) = \text{GP}(\rho_m, \text{cov}(t, t', \sigma, l)).$$

(12)

Here, $\rho_m$ is the mean density, defined as

$$\rho_m = \rho(l_1) = E[\rho].$$

(13)

For now, this is a temporally constant mean; if an experimentalist can make more specific assumptions on temporal trends, this will be easily included in further developments.
As covariance function, we use the so-called squared exponential kernel \( K_{\text{SE}} \) (Rasmussen, 2006) with time lag \( r = (t - t') \), variance \( \sigma_{\rho}^2 \), and correlation length \( l \), such that

\[
K_{\text{SE}}(r) = \sigma_{\rho}^2 \exp\left(-\frac{r^2}{2l^2}\right)
\]  

(14)

Expressing the mean as a constant and the covariance as a function of lag \( r = (t - t') \) makes the GP stationary, which means that all its statistics are invariant with respect to shifts in time. Among others, this means that there are no trends—neither in the mean nor in the expected amplitude of fluctuations, nor in their correlation. Choosing the squared exponential kernel leads to very smooth functions \( \rho(t) \), while, e.g., the plain exponential kernel would lead to continuous but non-differentiable \( \rho(t) \). For illustration, Fig. 12 (Appendix C) shows a collection of randomly generated functions from a GP.

2.6 Analytical Approach

With \( \rho(t) \) modeled as a GP, we can quantify the errors \( \sigma_{\text{noise}}^2 \) and \( \sigma_{\text{raster}}^2 \). The analytical approach is much faster because it avoids the Monte-Carlo simulation. It is based only on analytically available statistics of GPs and of the Poisson distribution, and uses linear error propagation from intensities to densities via a first-order Taylor expansion of Beer–Lambert’s law.

2.6.1 Raster Error

We define the raster error \( \sigma_{\text{raster}}^2 \) (or dispersion variance in the context of GPs):

\[
\sigma_{\text{raster}}^2 = E\left[\left(\rho(t) - \hat{\rho}(t)\right)^2\right] = \sigma_{\rho}^2 - \sigma_s^2,
\]

(15)

where \( \sigma_s^2 \) is the total variance of the time-varying density (i.e., the variance of the GP), and \( \sigma_{\rho}^2 \) is the so-called smoothed variance (i.e., the variance recovered by the rastered time series \( \hat{\rho}(t) \)). The term dispersion variance is a statistical term (where dispersion relates to variance), not to be confused with optical dispersion. The smoothed variance \( \sigma_s^2 \) can be quantified as

\[
\sigma_s^2 = \frac{1}{L^2} \int_0^L \int_0^L \text{cov}(t, t', L, \sigma_{\rho}) dt dt',
\]

(16)

where \( L \) is the exposure time. Combining these two equations, the raster error yields

\[
\sigma_{\text{raster}}^2 = \sigma_{\rho}^2 - \frac{1}{L^2} \int_0^L \int_0^L \text{cov}(t, t', L, \sigma_{\rho}) dt dt'.
\]

(17)

The derivation of the analytical solution \( F_{\text{SE}}(L) \) for the double integration in Eq. (17) for the squared exponential kernel is provided in Appendix A, Eq. (46):

\[
F_{\text{SE}}(L) = \int_0^L \int_0^L \text{cov}(t - t') dt dL_1 dL_2 = \sqrt{2\pi} \sigma_s^2 L_0 \cdot \text{erf}\left(\frac{L}{\sqrt{2}l}\right) + 2l^2 [\text{cov}(L) - \text{cov}(0)].
\]

(18)
Note that \( \text{cov}(r = 0) = \sigma^2 \). Therefore, the raster error follows as

\[
\sigma_{\text{raster}}^2 = \sigma_\rho^2 - \frac{1}{L^2} F_{\text{SE}}(L).
\]  

(19)

From Eqs. (18) and (19), the raster error \( \sigma_{\text{raster}}^2 \) depends, besides on the desired time window size \( L \), on the anticipated variance \( \sigma_\rho^2 \) of the variable density, its correlation length \( l \), and on the entire covariance function \( \text{cov}(\cdot) \) of the GP. These are the statistical expectations concerning the amplitude and frequencies of \( \tilde{\rho}(t) \). In the current context, they describe how difficult it is to approximate the true dynamics with a rastered curve \( \tilde{\rho}(t) \) that is piecewise constant per time window of length \( L \).

### 2.6.2 Noise Error

Recall that photon counts \( k_L(-) \) counted in a time window \( L(s) \) would be converted to a rate \( \kappa(1/s) \) via Eq. (6). Also, recall that, before conducting the experiment, we statistically anticipate results by random variables. Here, this means to work with the random count rate \( \tilde{\chi} \) rather than with the actual count rates \( \kappa \). The random count rates \( \tilde{\chi} \) have an expected value of \( I_1 \) given by Eq. (4) and Eq. (8): For any set of Beer–Lambert parameters \( [\varepsilon, \rho, d, I_0] \), we get an expected count number \( \lambda \) per unit time via Eq. (4), and then convert \( \lambda \) to the expected count rate \( \chi_0 = E[\tilde{\chi}] \). Then, according to Eq. (9), we can anticipate their imprecision as \( \sigma_{\chi}^2 = I_1/L \). Equation (4) in this procedure requires a value for density, for which we use the expected density \( \tilde{\rho}_m \) that we already used as part of the GP.

Yet, this is only the imprecision of the random count rate. What we need now is to translate it to the noise variance \( \sigma_{\text{noise}}^2 \) of the densities \( \rho \) that we would compute from the Beer–Lambert’s law with the point estimate \( I_1 \approx \kappa \). Thus, we need an uncertainty propagation from \( \chi \) (in lieu of \( I_1 \approx \kappa \)) via Beer–Lambert’s law (Eq. (1)) to densities \( \rho \), Since the Beer–Lambert equation \( \rho = \rho(\chi) \) is nonlinear, we use linearized error propagation via Taylor-expansion, truncated after the first derivative as announced above:

\[
\rho(\chi) \approx \rho(\chi_0) + \frac{\partial \rho}{\partial \chi} \bigg|_{\chi_0} (\chi - \chi_0),
\]

(20)

where we use \( \chi_0 = E[\tilde{\chi}] \) as expansion point. In Appendix B, we provide the derivation of the noise error from the above considerations and Eq. (20):

\[
\sigma_{\text{noise}}^2(L) = \frac{\exp(\varepsilon \rho_m d)}{I_0 L \varepsilon^2 d^2}.
\]

(21)

This expression contains all Beer–Lambert parameters that make up the attenuation, including the mean density \( \rho_m \) as statistical expectation, and is antiproportional to the time window \( L \), i.e., the noise error decreases with increasing \( L \).

### 2.7 Numerical Approach

In contrast to the first-order analytical approximation in Sect. 2.6, the numerical approach offers arbitrarily high accuracy to compute the total error \( \sigma^2_{\text{total}}(L) \), yet only at the limit of infinite computer time spent on the involved Monte-Carlo simulations. It randomly simulates the photon counter by randomly generating realizations of \( \rho(t) \) from the GP (Fig. 4a). Then it produces random numbers of counts \( k_L \) per time window of length \( L \) by randomly
drawing from the Poisson distribution (Eq. (2)). Upon conversion to count rates \( \kappa \), this yields random time series \( \kappa(t) \), used as a point estimate for \( \hat{I}_f(t) \). For comparison, the same is done without Poisson noise to receive a noise-free reference curve. Then, we convert both time series (noisy and idealized, noise-free count rates) to density raster curves \( \hat{\rho}_L(t) \) via Beer–Lambert’s law. Finally, we compare the noise-affected versions of the raster curve for density (Fig. 4c) with the noise-free raster curve (Fig. 4b) and with the originally simulated exact curve (Fig. 4a) to measure the errors.

When this procedure is repeated many times \( (n_{MC}) \) via Monte-Carlo simulation, we can numerically extract the total error as defined in Eq. (10), and also extract the individual terms for \( \sigma^2_{\text{raster}} \) (Fig. 4b) and \( \sigma^2_{\text{noise}} \) (Fig. 4c). For the latter, it is necessary to generate many \( \kappa(t) \) realizations for each realization of \( \hat{\rho}(t) \).

### 2.8 Optimization and Implementation

With \( \sigma^2_{\text{total}}(L) \) expressed analytically in Sect. 2.6, it can be seen that the optimization problem in Eq. (11) is convex and differentiable. Hence, we can find \( L_{\text{opt}} \) by setting the derivative to zero:

\[
L_{\text{opt}} = \frac{\partial \sigma^2_{\text{total}}(L)}{\partial L} = 0
\]

The derivative of the noise error (Eq. (21)) is

**Fig. 4** Steps within the numerical quantification of errors for one density sample: a Density realization \( \hat{\rho}(t) \) randomly generated from the GP; b Rastered density realization \( \hat{\rho}(t) \) computed noise-free from \( \hat{\rho}(t) \) to quantify the raster error against the original realization \( \hat{\rho}(t) \); c Many realizations of noise-affected step curves \( \hat{\rho}(t) \) for quantifying the noise error by comparison to the noise-free step curve from (b)
The derivative of the raster error (Eq. (49) in Appendix A) is

\[
\frac{\delta \sigma_{\text{raster}}^2(L)}{\delta L} = \frac{2}{L^3} F(L) - \frac{1}{L^2} F(L)'.
\]

Therefore, the derivative of the total error is

\[
\frac{\delta \sigma_{\text{noise}}^2(L)}{\delta L} = \frac{\delta \sigma_{\text{noise}}^2(L)}{\delta L} + \frac{\delta \sigma_{\text{raster}}^2(L)}{\delta L} = \frac{-1}{(\epsilon d)^2 E[I_1]} \frac{1}{L^2} + \frac{2}{L^3} F(L) - \frac{1}{L^2} F(L)',
\]

where \(F(L)\) and \(F(L)’\) are defined in Eqs. (46) and (47), respectively. While analytical solutions may be achievable for simpler covariance functions, for our current case we find the roots of Eq. (24) numerically. For this, we use the function fzero as implemented in MATLAB (The Mathworks, Inc., 2016).

For solving the optimization based on the numerical approach from Sect. 2.7, we call the numerical procedure described there for a finely resolved list of window sizes \(L\) (in steps of one second) and pick the optimal value \(L_{\text{opt}}\) manually. The next figure shows a flowchart that summarizes the steps to optimize the exposure time (Fig. 5).

---

**Fig. 5** Flowchart summarizing the steps to optimize the exposure time
2.9 Scenarios

To illustrate our suggested approach, we first use a base case scenario, whose parameters are included in Table 1. The incident intensity (count rate $I_0 = 400/\text{s}$) can be seen as a typical value from literature reports. It is chosen such that variations by factors of up to 0.2 and 5 as part of our later sensitivity analysis are still meaningful. In actual experiments, $I_0$ can be controlled by choice of the source material (e.g., Taqi & Khalil, 2017) and through technical details of the collimator (e.g., Mayar et al., 2019).

As the width of the system, we use a relatively thin width of $d = 2 \text{ cm}$. This is mainly to ensure that variations by the same factors as for $I_0$ still yield plausible values. As mean density $\rho_m = 2650 \text{ kg/m}^3$ and absorption concentration $\varepsilon = 70.471 \times 10^{-3} \text{ m}^2/\text{kg}$, we take values representative of aluminum as used in recent GRA experiments by Mayar et al. (2019). Aluminum, by pure chance, has the same density as quartz, which is the main constituent of many sand-based natural porous media.

For the time-dependent density, we take a GP variance $\sigma_r^2 = (300 \text{ kg/m}^3)^2$, i.e., a standard deviation of $\sigma_r = 300 \text{ kg/m}^3$. As GPs assume a Gaussian distribution, this implies that about 95% of all variable densities should fall into the interval of $\rho_m \pm 2 \cdot \sigma_r = [2050, 3250] \text{ kg/m}^3$. This is a proportion that could be caused by changes in the porosity of a porous material.

As correlation length for the GP, we take $l = 100 \text{ s}$. This choice is almost arbitrary since the density processes to be monitored could be—depending on the experimental conditions and the considered density-changing processes—between seconds and months for hydraulic-flume colmation (Mayar et al., 2020) or experimental investigations of mineral precipitation in porous media (Hommel et al., 2015). Again, the variations by factors of up to 0.2 and 5 are still plausible values.

We are not fixing a value for exposure time $L$, because the choice of exposure time will be subject to screening as part of the optimization.

To investigate sensitivity questions about the optimal experimental design, we also run a series of scenario variations. We change one parameter at a time: Beer–Lambert parameters ($I_0, d, \varepsilon, \rho_m$) and GP hyperparameters ($\sigma_r, l$). These scenarios are defined in Table 2. For all varied parameters, we divide and/or multiply their base values by factors of 5, 2, and 1.33. A few of the resulting parameter combinations may be atypical for realistic experimental conditions, but we prefer having systematic and identical variations across all parameters for clearer interpretation.

| **Table 1** Values of the parameters used in the base case scenario |
|-------------------|----------------|
| Parameter (units) | Base value |
| $I_0(\text{s}^{-1})$ | 400 |
| $d(\text{m})$ | 0.02 |
| $\rho_m(\text{kg/m}^3)^*$ | 2,650 |
| $\varepsilon(\text{m}^2/\text{kg})^{**}$ | $7.471 \times 10^{-3}$ |
| $\sigma_r(\text{kg/m}^3)$ | 300 |
| $l(\text{s})$ | 100 |

*Equivalent to quartz sand and aluminum
**Equivalent to aluminum
Table 2  Values of the parameters for sensitivity analysis. Other parameters are fixed (as in Table 1)

| Scenario series | Parameter | Values                   |
|-----------------|-----------|--------------------------|
|                 | \( \sigma_s (\text{kg/m}^3) \) | \( \frac{1}{5} \) | \( \frac{1}{2} \) | \( \frac{1}{133} \) | Base case | \( \times 1.33 \) | \( \times 2 \) | \( \times 5 \) |
| 1               | 60        | 150                      | 225                      | 300                      | 400       | 600       | 1500       |
| 2               | 20        | 50                       | 75                       | 100                      | 133.3     | 200       | 500        |
| 3               | 80        | 200                      | 300                      | 400                      | 533.3     | 800       | 2000       |
| 4               | 0.002     | 0.005                    | 0.0075                   | 0.01                     | 0.013     | 0.02      | 0.05       |
| 5               | \( \epsilon (\text{m}^2/\text{kg}) \) | \( 1.494 \times 10^{-3} \) | \( 3.736 \times 10^{-3} \) | \( 5.600 \times 10^{-3} \) | \( 7.471 \times 10^{-3} \) | \( 99.36 \times 10^{-3} \) | \( 149.42 \times 10^{-3} \) | \( 373.55 \times 10^{-3} \) |
| 6*              | \( \rho_m (\text{kg/m}^3) \) | 530                      | 1000                     | 1987.5                   | 2650      | 3533.3    | 5300       | 13,250     |
3 Results and Discussion

3.1 Base Case and Comparison of the Analytical and Numerical Approach

Figure 6 shows the results of the analytical and numerical approaches. It shows both the individual error terms (noise error: red lines, raster error: blue lines) and the resulting total error (black lines) over window size $L$. As the difference between analytical results (solid lines) and numerical results (dashed lines) is barely invisible in Fig. 6, we anticipate two scenario variations by using a width of $d = 0.01$ m and $d = 0.10$ m to confirm this. Since their plots show essentially the same behavior as Fig. 6, we show the corresponding results in Appendix C (Fig. 13).

At first in Fig. 6, the noise error decreases with $L$, while the raster error increases with $L$, as expected from Section 2.3. Both approaches, analytical and numerical, provide very similar results. Visible differences are observed in the raster error for $L$ greater than 80 s (Fig. 6).

The vertical cyan line in Fig. 6 shows $L_{\text{opt}}$ for the analytical approach, while the vertical magenta line shows $L_{\text{opt}}$ for the numerical approach. Both optimal solutions are very similar with $L_{\text{opt,analytical}} = 48.8$ s and $L_{\text{opt,numerical}} = 49$ s ($L$ in the numerical approach is scanned at a resolution of 1 s). Since these solutions are very similar and because the analytical approach is much faster (3 s versus 2 h on a regular, contemporary desktop computer), we will use the analytical approach for the scenarios variation (Table 2) and will recommend it for use in future studies.
3.2 Influences of Parameters on Optimal Window Size and Attainable Minimal Errors

We grouped the results into two groups: parameters that affect the raster error and parameters that influence the noise error. Looking at the equations of both errors (Eqs. (19) and (21), respectively), we can see that the raster error is influenced only by those hyperparameters of the GP that describe the changes of density ($\sigma_\rho^2$ and $l$). In contrast, the noise error is influenced only by the parameters of Beer–Lambert’s law ($I_0$, $\rho_m$, $d$, and $\varepsilon$).

3.2.1 Influence of Signal Variance and Correlation Length on the Raster Error

To analyze the sensitivity of $L_{\text{opt}}$ with respect to the expected amplitude of density (represented by the GP $\sigma_\rho^2$), we vary $\sigma_\rho^2$ according to the scenario variation 1 of Table 2. Note that results are shown as a function of the density’s standard deviation $\sigma_\rho$ and not the variance $\sigma_\rho^2$.

As Fig. 7a shows, when $\sigma_\rho$ is increased, the raster error increases too. The noise error is not affected, since it does not depend on $\sigma_\rho$ (Eq. (21)). The rise of raster error forces $L_{\text{opt}}$ to decrease when $\sigma_\rho$ is increased (Fig. 7a, b). In other words: more pronounced changes of $\rho(t)$ need a finer resolution (smaller $L_{\text{opt}}$), i.e., the stronger changes occur in a system, the more “snapshots” will be needed to capture the whole dynamic. This is intuitive and not surprising, but now we can quantify and produce an optimal experimental plan. Also in Fig. 7b, we observe that the optimal total error $\sigma_{\text{total}}^2 (L_{\text{opt}})$ decreases very strongly with decreasing strength of changes (i.e., with decreasing $\sigma_\rho^2$).

Similarly, to analyze the sensitivity of $L_{\text{opt}}$ with respect to the correlation length $l$, we vary $l$ according to scenario variation 2 in Table 2. We recall that the length scale $l$ of the GPE represents the characteristic time scale on which we anticipate the changes of $\rho(t)$ to act. As seen in Fig. 8a, when $l$ increases (i.e., the dynamics are slower), the raster error decreases; and consequently the total error decreases too. Again, the noise error is not affected since it does not depend on the $l$ (Eq. (21)). When $l$ increases, $L_{\text{opt}}$ increases too (Fig. 8b).

![Fig. 7](https://via.placeholder.com/150)

**Fig. 7** a Errors and b optimal window sizes and total error as a function of the density’s standard deviation in Table 2
This means that faster dynamics in $\rho(t)$ (smaller $l$) call for smaller $L$. Slower dynamics (larger $l$) allow for larger $L$, and hence for strongly reduced noise error while still resolving the dynamics well. Overall, the attainable total error is smaller for slower dynamics.

### 3.2.2 Impact of Incident Intensity, Thickness, Substance Absorption, and Mean Density on the Noise Error

To analyze the sensitivity of $L_{\text{opt}}$ with respect to the incident intensity $I_0$, we vary $I_0$ according to the scenario variations 3 in Table 2. As seen in Fig. 9a, when the experimentalist increases $I_0$, the noise error decreases (Eq. (21)), and consequently the total error decreases too. The raster error is not affected since it does not depend on $I_0$ (Eq. (19)). When $I_0$ increases and the importance of noise error fades, $L_{\text{opt}}$ can decrease (Fig. 9b) to better resolve the changes at smaller total errors. Therefore, the clear recommendation is to use a strong source. If the source cannot be varied, one can also use a larger collimator to obtain larger $I_0$. Using larger collimators, however, would reduce the spatial resolution if density were spatially heterogeneous. For such cases, the optimal collimator size could be determined as shown by Mayar et al. (2019).

Parameters $d$ and $\varepsilon$ have identical roles in Eq. (24) via Eq. (4), therefore results should be identical. For checking this, we provide both sets of results. That means, we vary $d$ according to scenario 4 in Table 2; additionally, we proceed with $\varepsilon$ according to scenario 5 of Table 2. Results for the impact of $d$ and $\varepsilon$ are displayed in Fig. 10.

We observe in Fig. 10a, b that only the noise error (Eq. (21)) is affected by $d$ and $\varepsilon$, respectively; the raster error is not affected as already visible from the respective equations. The noise error decreases when increasing the thickness $d$, at least for the values studied in Table 2. $L_{\text{opt}}$ also decreases when decreasing $d$. However, this decreasing behavior changes when we apply a broader range of thicknesses. These results are shown and discussed in Appendix C (Fig. 14).

The impact of $\varepsilon$ (Fig. 10c, d) shows the same non-monotonic behavior as $d$ on the $L_{\text{opt}}$. In fact, except for a change of scale on the x-axis, all results are identical in theory and practice.
Finally, we analyze the sensitivity of $L_{\text{opt}}$ with respect to $\rho_m$. That means we vary $\rho_m$ according to scenario 6 of Table 2. We observe in Fig. 11a that, when the mean density increases, the noise error is also increased, while the raster error remains constant (as expected from the respective equations). When $\rho_m$ increases the noise error, $L_{\text{opt}}$ has to increase as well (Fig. 11b) to compensate for the rising noise by longer exposure. However, the observable increments of the total error or the change in $L_{\text{opt}}$ for the studied range of $\rho_m$ in Table 2 are by far not as large as for the other parameters.

Comparing the behavior of $\rho_m$ with that of $d$ and $\varepsilon$, one may wonder about the differences, as $\rho_m$ seems to have the same role in Beer–Lambert’s law (Eq. (1)) as $d$ and $\varepsilon$. However, $\rho_m$ is mathematically different, as it is the quantity of interest we want to observe. Instead of occurring quadratically in the denominator of Eq. (21), it shows up in (expected) quadratic form on the left-hand side as noise variance. From that point of view, it could be moved into the denominator in (expected) quadratic form just as well.

### 3.3 Overall Optimal Experimental Design

Recall that both $d$ and $\varepsilon$ show a non-monotonic influence on $L_{\text{opt}}$ and on the attainable total error. Instead, they show a convex impact, just like the exposure time $L$. Convex is the mathematical term that refers to an upward-curved function, and guarantees to have a well-defined and unique minimum. Altogether, the total error is a convex function of $L, d,$ and $\varepsilon$, and hence will possess an overall joint optimum with respect to these three parameters. Figure 14 (in Appendix C) shows individual cross-sections through this function. That figure visualizes, at least, pairwise optima. The overall optimum among all three at the same time can be even better.

The location of this combined optimum (across $L, d,$ and $\varepsilon$) depends on all remaining parameters considered in this study ($I_0, \rho_m, \sigma_p^2, l$). All four of these showed monotonous behavior—hence extreme values of them are the best for the experiment. However, not all of them are easy to control at extreme values. Incident intensities have a manageable upper limit, and the remaining three parameters $\rho_m, \sigma_p^2$ and $l$ are given by the density process one wishes to monitor.

Of course, the experimentalist can try to stimulate the time-dependent density process to be strongly pronounced but slow for a best-possible signal. This facilitates an overall iteration between technical optimization (for $L, d,$ and $\varepsilon$) and other experimental controls (affecting $\rho_m, \sigma_p^2, l$): First, choose the largest available $I_0$, fix the statistical expectations $\rho_m, \sigma_p^2$ and $l$, and use the $\varepsilon$ of the planned material. Then, take our expressions for total error (Eqs. (10), (19), and (21)), and optimize jointly across $L$ and $d$. If the material itself is also exchangeable, include $\varepsilon$ within this optimization. If the overall attainable, minimal errors are not satisfying, think about process controls to modify $\rho_m, \sigma_p^2$ and $l$, until finding a promising overall setup.

### 3.4 Limitations

Working with statistical expectations is the best one can do before conducting an experiment, and our approach relies on GPs as a tool for doing this. Yet, the anticipation of time-dependent densities by a GP may be a difficult guess in practice. In addition, the real density curve to result from the planned experiment may be more complex than the realization of a stationary GP. For example, it may exhibit non-stationary in the mean (i.e., trends) or in its dynamics (i.e., changing from slow to fast), or it may exhibit abrupt jumps. For such
Fig. 9  a Errors and b optimal window sizes and total error as a function of incident intensity in Table 2

Fig. 10  a Errors and b optimal window sizes and total error as a function of thickness in Table 2. c Errors and d optimal window sizes and total error as a function of absorption of the substance $\epsilon$ in Table 2
cases, more detailed and involved modeling than a stationary GP could be advised. We chose GPs, because they are mathematically convenient to handle, and because we are convinced that the key parameters of a GP represent pieces of information that an experimentalist may actually be able to foresee with some confidence. However, the field of stochastic processes and random functions is much richer than just GPs.

A second limitation is that we allow density to be time-dependent, but ignore that it may be variable in space. In spatially heterogeneous systems, there will be an additional need for spatial scanning (grid resolution), on top of the question of temporal rastering (sampling rate). Then, the question investigated here will expand: how long to record photon counts at each spatial location before moving on to the next position, and how fine to resolve spatially between these different positions? Together, the time per location and the number of locations provide a turnaround time before revisiting locations. Then, one would desire a minimum total error that will consist of spatial raster error, temporal raster error, and noise error. The contribution in the current paper is the foundation for such further research.

A related limitation of the sensitivity analysis (not of our optimization method) affects the statement that decreasing the thickness also decreases $L_{opt}$ and will lead to lower overall errors. This is correct as long as there is a reliable density of the material. Under realistic conditions, e.g., when analyzing the density of porous media, the (bulk) density of thin samples is not yet statistically stable. This will hold for all sample dimensions that are still below the size of a representative elementary volume. Then, changing the thickness may also change the bulk density in a hard-to-predict manner, destroying the independence between density and thickness. In our theoretical study, we assumed these two factors to be independent. In practice, the corresponding aspects of the sensitivity study will hold on average but are not guaranteed to hold for each individual case.

Finally, recall that we assumed a monochromatic source of gamma rays, so that the description by Beer–Lambert’s law is accurate. However, applying our framework to polychromatic GRA experiments is a good approximation when the absorption coefficient does not change strongly across the involved wavelengths. Otherwise, the Taylor expansion for linearized error propagation from intensities to densities needs to be written as an average.

Fig. 11  a Errors and b optimal window sizes and total error as function of mean density in Table 2
of Taylor expansions across the polychromatic spectrum. This is possible, but would simply inflate the notation.

4 Summary and Conclusions

In this study, we mathematically modeled and then minimized errors that arise when monitoring the density (caused by volume fractions of individual constituents in multiphasic/porous media) as a function of time by GRA experiments. In such experiments, one has to choose a so-called exposure time, which is the time spent on counting photon impacts of the attenuated gamma rays after passing through the systems, before converting the count to an intensity and then to a density value. We showed that the total error in observing time-dependent densities has two terms: the raster error and the noise error. Raster error arises by approximating the true variable density by a raster curve, consisting of piecewise constant density values per exposure time window. The noise error arises due to the stochastic nature of gamma-ray emission and attenuation. Both errors are controlled by the exposure time, but in contradicting directions: long exposure times capture the dynamics poorly (raster error), whereas short exposure times yield imprecise values (noise error). These two errors sum up and pose competing influences on the optimal choice of exposure times.

We provided an optimization framework for experimental designs in dynamic GRA measurements. As per deviation, it minimizes the total error by providing the optimal exposure time, but it can serve just as well to optimize other key experimental parameters that are part of Beer–Lambert’s law. We modeled the yet unknown dynamics (i.e., mean and amplitude of density and characteristic time of density changes) as a random Gaussian process. We applied an analytical approach by a first-order approximation of total errors and we compare it to a numerical approach (Monte-Carlo simulation). The latter is arbitrarily accurate at the limit of infinite computing time, yet both approaches produce similar results. Thus, we recommend and used the analytical approach since it is very fast and is simpler to implement.

We demonstrated our approach on a synthetic experimental case and discussed the composition of total error, the attainable minimum of errors, and the optimal exposure time. The optimum still depends on other experimental parameters that occur in the Beer–Lambert attenuation law (incident intensity, mean density, thickness, and substance absorption). It also depends on the statistical expectations on the changes of density expressed through parameters of the GP. The latter include the variance (expected amplitude of density changes) and the correlation length (characteristic time of density changes) of the GP. To capture their influence, we conducted a sensitivity analysis by systematic scenario variation and discussed the results.

Our results showed that the signal variance and correlation length influence only the raster error, while the Lambert–Beer parameters influence only the noise error. However, the impact of the mean density (within the range studied here) is low. Faster and stronger dynamics of the process (large signal variance and small correlation length) increase the raster error and hence require smaller exposure times to capture these strong and fast dynamics. The noise error decreases with increasing incident intensity, allowing for smaller exposure times to better resolve the dynamics at smaller total errors.

Thickness and substance absorption jointly showed a special, non-monotonic behavior on the optimal exposure time when their values are changed. Both parameters share
two competing effects: larger values increase the strength of the observable dynamics in intensities, but larger values also lead to an exponential decrease of output intensities. For small values, the first effect dominates, so that increasing these parameters leads to smaller optimal exposure times and smaller total errors. Beyond a specific break-even point, both the total error and the optimal exposure time rise again. This break-even point presents an overall optimum of GRA experiments. Therefore, we also illustrated and discussed how to apply our optimization to a joint optimization of exposure time, thickness, absorption coefficient and process controls that affect the time-variability of density.

Overall, this study offers a quantitative knowledge of inevitable errors to experimentalists and ways to minimize them. Our recommendations for the experimentalist are the following: to choose larger exposure times for slower dynamics, choose smaller exposure times for faster dynamics, use a strong source of intensity whenever possible, and consider the more complex interplay of the expected density and absorption coefficient of the material at study.

In future research, our model of errors can be extended by a spatial rastering error, and then an optimization can be found for temporally and spatially rastered monitoring of densities.

Appendix A: Derivation of analytical raster error

Integration of exp(·) based on covariance functions

Let us define the squared exponential kernel (the covariance function used) as function $K: \mathbb{R} \rightarrow \mathbb{R}$ by

$$K(x) := \sigma^2 \exp\left(-\frac{x^2}{\lambda}\right) \text{for} \lambda, \sigma > 0. \quad (25)$$

First, let us recapitulate some basic properties of $K(·)$, which we will use later. Due to the properties of exp(·) we have

$$K'(x) = -2\frac{x}{\lambda} K(x), \quad (26)$$

$$\partial_x K(x - y) = -\partial_y K(x - y), \quad (27)$$

$$\int_a^b x K(x) dx = -\frac{1}{2} \left[ K(x) \right]_a^b, \text{by Eq. (26)}. \quad (28)$$

Further, we use the following definition of the erf(·) function

$$\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \exp(-\tau) d\tau. \quad (29)$$

Now let us consider the integral

$$\int_0^{L_1} \int_0^{L_2} K(x - y) dy dx. \quad (30)$$
Using integration by parts and Fubini-theorem, we obtain by using Eq. (27), (28).

\[
\int_0^{L_2} \int_0^{k(x-y)dx} = \int_0^{L_1} \int_0^{k(x-y)dy} = \int_0^{L_1} \left( K(x-y) \right) dy \int_0^{L_2} \left( K(x-y) \right) dx - \int_0^{L_2} \int_0^{k(x-y)dy} = \int_0^{L_1} \left( K(x-y) \right) dy \int_0^{L_2} \left( K(x-y) \right) dx
\]

Considering \( I = \int_0^{L_2} yK(L_1 - y)dy \) and by using Eq. (28) we obtain

\[
I = \int_0^{L_2} yK(L_1 - y)dy = \int_0^{L_2} L_1K(L_1 - y)dy - \frac{\partial}{2} \left[ K(L_1 - L_2) - K(L_1) \right] + L_1 \int_0^{L_2} K(L_1 - y)dy
\]

This yields

\[
\int_0^{L_2} K(x-y) dx = \int_0^{L_2} L_2K(x-L_2)dx + \int_0^{L_2} L_1(y - L_1)dy + \frac{\partial}{2} \left[ K(L_1) + K(L_2) - K(L_1 - L_2) - K(0) \right]
\]

Using integration by substitution and Eq. (25) we obtain for \( a, b > 0 \)

\[
\int_0^a K(x-b)dx = \sigma^2 \int_0^a \exp \left( -\frac{(x-b)^2}{\lambda} \right) dx = \sigma^2 \sqrt{\lambda} \int_0^a \exp \left( -\frac{(x-b)^2}{\lambda} \right) \frac{1}{\sqrt{\lambda}} dx = \sigma^2 \sqrt{\lambda} \int_{-b/\sqrt{\lambda}}^{a-b/\sqrt{\lambda}} \exp(-r^2)dr
\]

Combining Eqs. (33), (34), and (29) we obtain

\[
\int_0^{L_1} \int_0^{L_2} K(x-y) dy dx = \sigma^2 \sqrt{\lambda} \int_{-L_2/\sqrt{\lambda}}^{L_1} \exp(-r^2) dr + \frac{\partial}{2} \left[ K(L_1) + K(L_2) - K(L_1 - L_2) - K(0) \right]
\]

\[
= \sigma^2 \sqrt{\lambda} \int_{-L_2/\sqrt{\lambda}}^{L_1} \exp(-r^2) dr + \frac{\partial}{2} \left[ K(L_1) + K(L_2) - K(L_1 - L_2) - K(0) \right]
\]

\[
= \frac{\sqrt{\pi}}{2} \left[ \frac{1}{2} \left( \text{erf} \left( \frac{L_1 - L_2}{\sqrt{\lambda}} \right) + \text{erf} \left( \frac{L_2}{\sqrt{\lambda}} \right) \right) + L_1 \left( \text{erf} \left( \frac{L_2 - L_1}{\sqrt{\lambda}} \right) + \text{erf} \left( \frac{L_1}{\sqrt{\lambda}} \right) \right) \right]
\]

Summarized, we obtain

\[
\int_0^{L_1} \int_0^{L_2} K(x-y) dy dx = \frac{\sqrt{\pi}}{2} \left[ \frac{1}{2} \left( \text{erf} \left( \frac{L_1 - L_2}{\sqrt{\lambda}} \right) + \text{erf} \left( \frac{L_2}{\sqrt{\lambda}} \right) \right) + L_1 \left( \text{erf} \left( \frac{L_2 - L_1}{\sqrt{\lambda}} \right) + \text{erf} \left( \frac{L_1}{\sqrt{\lambda}} \right) \right) \right] + \frac{\partial}{2} \left[ K(L_1) + K(L_2) - K(L_1 - L_2) - K(0) \right]
\]

**Derivatives**

Let us define
Here we use the squared exponential kernel with variance $\sigma^2$ and correlation length $l$: 

$$K_{SE}(r) = \sigma^2 \exp\left(-\frac{r^2}{2l^2}\right)$$  

(44)  

For two different upper bounds $L_1$ and $L_2$, using Eq. (36) with $\lambda = 2l^2$ we obtain 

$$F_{SE}(L_1, L_2) = \int_0^{L_1} \int_0^{L_2} K_{SE}(x-y) \, dy \, dx = \frac{\sqrt{\pi} \sigma^2}{\sqrt{2}} \left( L_2 \left( \text{erf}\left( \frac{L_1 - L_2}{\sqrt{\lambda}} \right) + \text{erf}\left( \frac{L_2}{\sqrt{\lambda}} \right) \right) + L_1 \left( \text{erf}\left( \frac{L_2}{\sqrt{\lambda}} \right) + \text{erf}\left( \frac{L_1}{\sqrt{\lambda}} \right) \right) \right)$$  

(45)
For the special case of the same upper bound $L_1 = L_2 = L$ we have

$$F_{SE}(L) = \int_0^L \int_0^L K_{SE}(x - y)dydx = \sqrt{2\pi} \sigma^2 I L \cdot \text{erf}\left(\frac{L}{\sqrt{2}I}\right) + 2I^2[K(L) - K(0)] \quad (46)$$

Therefore, for the derivatives for the same upper bound, we obtain using Eq. (40)

$$F_{SE}(L)' = \sigma^2\left(\sqrt{2\pi} I \cdot \text{erf}\left(\frac{L}{\sqrt{2}I}\right)\right) + 2L \cdot \exp\left(-\frac{L^2}{2I^2}\right) - 2LK_{SE}(L). \quad (47)$$

For sake of completeness, for some constant $c > 0$ and function $G(L) = c - \frac{1}{L^2} F(L)$

$$G(L)' = \frac{2}{L^3} F(L) - \frac{1}{L^2} F(L)' = \frac{1}{L^2} \left(\frac{2}{L} F(L) - F(L) '\right). \quad (48)$$

### Appendix B: Derivation of analytical noise error

First, we re-arrange the Beer–Lambert’s law for density $\rho$, and replace $I \approx \kappa$ by the random variable $\chi$ as an act of statistical anticipation:

$$\rho(\chi) = \frac{\log \left( \frac{I_0}{\chi} \right)}{\epsilon d} \quad (50)$$

Then, the announced Taylor expansion about the point $\chi_o = E[\chi] = I_1$ results in

$$\rho(\chi) = \frac{\log \left( \frac{I_0}{I_1} \right)}{\epsilon d} - \frac{1}{\epsilon d I_1} (\chi - I_1) \quad (51)$$

Next, we define the linearized relationship $\rho = a\chi + b$ and find by comparison of coefficients that

$$a = -\frac{1}{\epsilon d I_1}$$

$$b = \frac{1}{\epsilon d} \left( \log \left( \frac{I_0}{I_1} \right) + 1 \right) \quad (52)$$

Applying standard linear(ized) error propagation with $\text{Var}[a\chi + b] = a^2 \cdot \text{Var}[\chi]$ yields for the noise error $\sigma_{\text{noise}}^2$ is

$$\sigma_{\text{noise}}^2 \approx a^2 \text{Var}(\kappa) = a^2 \sigma^2_{\chi} = a^2 \frac{\sigma_{\text{error}}^2}{L} = a^2 \frac{I_1}{L} = \frac{1}{(\epsilon d I_1)^2} \frac{I_1}{L} \quad (53)$$
Simplifying terms leads to

\[ \sigma_{\text{noise}}^2(L) = \frac{1}{(\varepsilon d)^2 I_1 L} \] (54)

This still depends on \( I_1 \), for which we need an estimation. For this, we re-arrange the Beer–Lambert's equation (Eq. (4)) for \( I_1 \), and insert \( \rho_m \) as statistical expectation of the density. Thus, we get:

\[ \sigma_{\text{noise}}^2(L) = \frac{\exp(\varepsilon \rho_m d)}{(\varepsilon d)^2 I_0 L} \] (55)

### Appendix C: Supplemental results

For illustration, Fig. 12 shows a collection of randomly generated functions from a GP with squared exponential covariance, using different values of variance \( \sigma^2 \) and correlation length \( l \). The black point markers in the figures are points where the true values of \( \rho(t) \) are known precisely. This will not occur during GRA experiments due to noise error, but it serves well to illustrate the properties of GPs. One can see how all randomly generated lines go through these points, and then start spreading when moving away from the points. The rate of spreading (and the frequency of fluctuations after that) depends on the correlation length

![Fig. 12](image-url)  
**Fig. 12** Randomly generated functions from a GP with a squared exponential covariance: a large variance \( \sigma^2 \) and large correlation length \( l \), b small variance \( \sigma^2 \) and large correlation length \( l \), c large variance \( \sigma^2 \) and small correlation length \( l \), d small variance \( \sigma^2 \) and small correlation length \( l \)
l, while the overall width of spreading at sufficient distance from known points is governed by the variance parameter $\sigma^2$.

Figure 13 shows the results of the analytical and numerical approaches for several thicknesses. For the scenarios with a thickness of $d = 0.01$ m and $d = 0.10$ m, the optimal window size for the analytical and numerical approaches are the same: 74 s and 28 s, respectively.

As seen in Fig. 14a, the total error presents a minimum value at $d = 0.10$ m for the base case scenario (in Table 1). Correspondingly, the optimal value of the window size visible in Fig. 13c also presents a minimum for $d = 0.10$ m at $L_{opt} = 28$ s. Therefore, when the thickness is either decreased or increased with respect to $d = 0.10$ m, the optimal window size increases (Fig. 13c). What we see here is a competition between two effects. First, a larger width amplifies the effect of density, so that density is easier to measure. However, second, a larger width also decreases overall the countable photons that manage to pass through the system. This competition is the reason why $d$ shows up quadratically in the denominator of the noise variance in Eq. (21) (first effect) and exponentially in the enumerator of the same equation (second effect).

For small values of $d$, the first (quadratic) effect dominates. In that range, increasing $d$ improves the total error, and allows for smaller exposure times. Then, there is a break-even point at a specific value of thickness (here: $d = 0.10$ m). Beyond that point, i.e., for large values of $d$, the second (exponential) effect dominates. There, further increase of $d$ is counterproductive. Overall, this results in a non-monotonic (but convex) impact of thickness $d$ both on the optimal exposure time and on the attainable total error. The impact of $\varepsilon$ (Fig. 14b) shows the same non-monotonic, convex behavior with break-even point, for the same mathematical and physical reasons as $d$.

Appendix D: Distinction of densities

We are interested to determine a length $L$ of time window that is large enough to distinguish between two different densities $\rho_1$ and $\rho_2$ in GRA experiments, still ignoring that $\rho_1$ and $\rho_2$ may be part of a time-variable density. For this purpose, we write the required length $L = NL_1$, i.e., as $N$ times the unit length $L_1$. Hence, $N$ stands for the number of seconds. In this setting, we can estimate the expected count per unit time window $\lambda = E[X]$.
via the estimate \( \hat{\lambda}_i = \frac{X_i}{N} \), where the subscript \( i (= 1, 2) \) enumerates the two different densities and corresponding GRA experiments over a length \( L = NL_i \), and \( X_i \) are the counts observed in these two experiments over the duration \( L \).

For \( N > N_{\text{min}} = 30 \), the estimates \( \hat{\lambda}_i \) are sufficiently close to normally distributed, so we can use a typical \( z \)-test of the hypotheses \( H_0 : \lambda_1 = \lambda_2 \) versus \( H_1 : \lambda_1 \neq \lambda_2 \). The corresponding test statistic is:

\[
z_{\text{test}} = \frac{\hat{\lambda}_2 - \hat{\lambda}_1}{\sqrt{\frac{\hat{\lambda}_1^2}{N_1} + \frac{\hat{\lambda}_2^2}{N_2}}}
\]  

In a two-sided test, for a given significance level \( \alpha \), the critical value of \( z \) is:

\[
z_{\text{crit}} = z_{\alpha/2} = \Phi^{-1}(1 - \alpha/2),
\]

where \( \Phi^{-1}(\cdot) \) is the inverse of the cumulative distribution function for the normal distribution. The power \( \beta \) of the two-sided hypothesis test at level \( \alpha \) is given by

\[
\beta = \Phi \left( \frac{\frac{\lambda_2 - \lambda_1}{\sqrt{\frac{\hat{\lambda}_1^2}{N_1} + \frac{\hat{\lambda}_2^2}{N_2}}} - z_{\alpha/2}}{2} \right).
\]

Note that the significance level \( \alpha \) controls the ‘false-negative’ probability that one falsely rejects \( H_0 \) although it is true (i.e., one would deem two identical density values to be different due to the random imprecision of photon counts). The power \( \beta \) controls the ‘true-negative’ probability that one correctly rejects \( H_0 \) when it is not true (i.e., one correctly distinguishes the two different densities).
In our current context, this means we formulate a requirement for \( \beta \): our goal is to determine the number of seconds \( N \) of exposure time, so that the probability of correctly distinguishing two different densities \( (\rho_1, \rho_2) \) is greater than \( \beta_{\text{crit}} \). Without loss of generality, we assume \( \lambda_2 > \lambda_1 \). Thus, we have:

\[
\Phi \left( \frac{\lambda_2 - \lambda_1}{\sqrt{\frac{\lambda_1 + \lambda_2}{N}}} - \frac{z_{a/2}}{\lambda_2 - \lambda_1} \right) > p_{\text{crit}} \Rightarrow N > \frac{(\Phi^{-1}(p_{\text{crit}}) + \Phi^{-1}(1 - a/2))^2}{\lambda_2 - \lambda_1}. \tag{59}
\]

From Eq. (4) we know that \( \lambda_i = \frac{L_i I_0}{\exp(\epsilon d \rho_i)} \). We can insert this to obtain:

\[
N > \frac{(\Phi^{-1}(p_{\text{crit}}) + z_{a/2})^2}{L_1 I_0 (\exp(-\epsilon d \rho_2) - \exp(-\epsilon d \rho_1))}. \tag{60}
\]

Apparently, the required number of seconds depends not only on a difference of densities, but also on their overall magnitude, and on all other Beer–Lambert parameters that control the overall attenuation \( (\epsilon, d) \). Also, it depends on the parameters specifying the statistical reliability \( (a, \beta) \).

In Fig. 15, we illustrate the resulting relation as a function of density difference \( \rho_2 - \rho_1 \) for a set of values for \( \rho_1 \) taken from our table of scenarios (Table 2), and using the base values for \( \epsilon \) and \( d \) as in Table 1. Also, we arbitrarily choose typical standard values of \( a = 5\% \) and \( \beta = 95\% \).

In Fig. 15, one can clearly see: (1) how overall larger densities require longer exposure time; (2) how smaller density differences require exponentially more exposure time (for which reason we exclude density differences smaller than a threshold \( \Delta \rho_{\text{min}} \) from the plot); and (3) values of \( N \) smaller than 30 (\( = N_{\text{crit}} \)) should not be taken as exact numbers; yet density differences larger than 10 or 20 kg/m\(^3\) can be safely identified within 30 s under the shown conditions.

Fig. 15 Relationship between \( N_{\text{crit}} \) and \( \rho_2 - \rho_1 \) for a set of \( \rho_1 \) values
Acknowledgements This work was supported by the Collaborative Research Center 1253 CAMPOS, funded by the German Research Foundation (DFG, Grant Agreement SFB 1253/1 2017). Further support by the German Research Foundation (DFG) through Project Number 327154368—SFB 1313, Project Number 432343452, and through Germany’s Excellence Strategy—EXC 2075—390740016 (SimTech).

Author contributions Ana Gonzalez-Nicolas, Ilja Kröker, Assem Mayar, and Wolfgang Nowak, contributed to the study conception and design. Material preparation, data collection and analysis were performed by Deborah Bilgic, Ana Gonzalez-Nicolas, and Ilja Kröker. The first draft of the manuscript was written by Ana Gonzalez-Nicolas and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding Open Access funding enabled and organized by Projekt DEAL. This work was supported by the Collaborative Research Center 1253 CAMPOS, funded by the German Research Foundation (DFG, Grant Agreement SFB 1253/1 2017). Further support by the German Research Foundation (DFG) through Project Number 327154368—SFB 1313, Project Number 432343452, and through Germany’s Excellence Strategy—EXC 2075—390740016 (SimTech).

Data Availability The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Conflict of interests The authors have no relevant financial or non-financial interests to disclose.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/

References

Barth, G.R., Illangasekare, T.H., Rajaram, H.: The effect of entrapped nonaqueous phase liquids on tracer transport in heterogeneous porous media: Laboratory experiments at the intermediate scale. J. Contam. Hydrol. 67(1), 247–268 (2003). https://doi.org/10.1016/S0169-7722(03)00066-4
Baytaş, A. F., Akbal, S.: Determination of soil parameters by gamma-ray transmission. Radiat. Measur. 35(1), 17–21 (2002). http://www.sciencedirect.com/science/article/pii/S1350448701002530
Beckers, F., Haun, S., Noack, M.: Experimental investigation of reservoir sediments. E3S Web Conf., 40, 03030 (2018). doi:https://doi.org/10.1051/e3sconf/20184003030
Berger, M. J., Hubbell, J.: XCOM: Photon cross sections on a personal computer (NBSIR–87-3597). National Bureau of Standards. (1987). http://www.iaea.org/inis/collection/NCLCollectionStore/_Public/19/009/19009871.pdf?r=1
Buczyk, B.: Poisson distribution of radioactive decay, p. 4. MIT Department of Physics, Cambridge (2009)
Eckberg, D.K., Sunada, D.K.: Nonsteady three-phase immiscible fluid distribution in porous media. Water Resour. Res. 20(12), 1891–1897 (1984). https://doi.org/10.1029/WR020i012p01891
Ferguson, H., Gardner, W.H.: Water content measurement in soil columns by gamma ray absorption. Soil Sci. Soc. Am. J. 26(1), 11–14 (1962). https://doi.org/10.2136/sssaaj1962.03615995002600010004x
Gharbi, D., Bertin, H., Omari, A.: Use of a gamma ray attenuation technique to study colloid deposition in porous media. Exp. Fluids 37(5), 665–672 (2004)
Haight, F.A.: Handbook of the Poisson distribution. Wiley, Hoboken (1967)
Hommel, J., Lauchnor, E., Phillips, A., Gerlach, R., Cunningham, A.B., Helmig, R., Ebigbo, A., Class, H.: A revised model for microbially induced calcite precipitation: improvements and new insights based on recent experiments. Water Resour. Res. 51(5), 3695–3715 (2015). https://doi.org/10.1002/2014WR016503
Illangasekare, T.H., Ramsey, J.L., Jensen, K.H., Butts, M.B.: Experimental study of movement and distribution of dense organic contaminants in heterogeneous aquifers. J. Contam. Hydrol. 20(1), 1–25 (1995). https://doi.org/10.1016/0169-7722(95)00045-W

Lykos, P.: The Beer–Lambert law revisited: a development without calculus. J. Chem. Educ. 69(9), 730 (1992). https://doi.org/10.1021/ed069p730

Matheron, G.: Matheron’s theory of regionalised variables. In: V. Pawlowsky-Glahn & J. Serra, (Eds.). Oxford University Press, Oxford. (2019). doi:https://doi.org/10.1093/oso/9780198835660.001.0001

Mayar, M.A., Schmid, G., Wieprecht, S., Noack, M.: Optimizing vertical profile measurements setup of gamma ray attenuation. Radiat. Phys. Chem. 164, 108376 (2019). https://doi.org/10.1016/j.radphyschem.2019.108376

Mayar, M.A., Schmid, G., Wieprecht, S., Noack, M.: Proof-of-concept for nonintrusive and undisturbed measurement of sediment infiltration masses using gamma-ray attenuation. J. Hydraul. Eng. 146(5), 04020032 (2020). https://doi.org/10.1061/(ASCE)HY.1943-7900.0001734

Mayar, M.A., Haun, S., Schmid, G., Wieprecht, S., Noack, M.: Measuring vertical distribution and dynamic development of sediment infiltration under laboratory conditions. J. Hydraul. Eng. (2022). https://doi.org/10.1061/(ASCE)HY.1943-001980

Moreira, A. C., Filho, O. P., Cavalcante, F. H. M., Appoloni, C. R.: Determination of hydraulic conductivity of undisturbed soil column: a measurement accomplished with the gamma ray transmission technique. In: O. Dikinya (Ed.), Developments in hydraulic conductivity research. IntechOpen. (2011). doi:https://doi.org/10.5772/15866

Nofziger, D.L.: Errors in gamma-ray measurements of water content and bulk density in nonuniform soils. Soil Sci. Soc. Am. J. 42(6), 845–850 (1978). https://doi.org/10.2136/sssaj1978.03615995004200060001x

Oostrom, M., Dane, J.H., Wietema, T.W.: A review of multidimensional, multifluid, intermediate-scale experiments: flow behavior, saturation imaging, and tracer detection and quantification. Vadose Zone J. 6(3), 610–637 (2007). https://doi.org/10.2136/vzj2006.0178

Pankow, J. F., Cherry, J. A.: Dense chlorinated solvents and other DNAPLs in groundwater: History, behavior, and remediation. Waterloo Press. (1996). https://scholar.google.com/scholar_lookup?title=Dense+chlorinated+solvents+and+other+DNAPLs+in+groundwater%3A+history%2C+behavior%2C+and+remediation&author=Pankow%2C+J.+F.&publication_year=1996

Rasmussen, C. E.: G1. Rasmussen, CE & Williams, CKI Gaussian processes for machine learning. MIT Press. Cambridge, MA, USA, 38, 715–719 (2006)

Sail, Y., Marot, D., Sibille, L., Alexis, A.: Suffusion tests on cohesionless granular matter: experimental study. Eur. J. Environ. Civil Eng. 15(5), 799–817 (2011)

Schälchli, U., Abegg, J., & Hunzinger, L. (2002). Kolmation–Methoden zur Erkennung und Bewertung. Eidg. Anstalt Für Wasserversorgung, Abwasserreinigung Und Gewässerschutz EAWAG, Dübendorf, Schweiz.

Taqi, A.H., Khalil, H.J.: An investigation on gamma attenuation of soil and oil-soil samples. J. Radiat. Res. Appl. Sci. 10(3), 252–261 (2017). https://doi.org/10.1016/j.jrras.2017.05.008

The Mathworks, Inc. (2016). MATLAB (MATLAB version 9.1.0.441655 (R2016b)) [Computer software]. The Mathworks, Inc.

van Genuchten, M.T.: A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. Soil Sci. Soc. Am. J. 44(5), 892–898 (1980). https://doi.org/10.2136/sssaj1980.03615995004400050002x

Waggoner, M.A.: Radioactive decay of Cs 137. Phys. Rev. 82(6), 906–909 (1951). https://doi.org/10.1103/PhysRev.82.906

Werth, C.J., Zhang, C., Brusseau, M.L., Oostrom, M., Baumann, T.: A review of non-invasive imaging methods and applications in contaminant hydrogeology research. J. Contam. Hydrol. 113(1), 1–24 (2010). https://doi.org/10.1016/j.jconhyd.2010.01.001

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.
Authors and Affiliations

Ana Gonzalez-Nicolas\textsuperscript{1} · Deborah Bilgic\textsuperscript{1} · Ilja Kröker\textsuperscript{1} · Assem Mayar\textsuperscript{1} · Luca Trevisan\textsuperscript{2} · Holger Steeb\textsuperscript{3} · Silke Wieprecht\textsuperscript{1} · Wolfgang Nowak\textsuperscript{1,16}

Deborah Bilgic
Deborah.Bilgic@de.bosch.com
Ilja Kröker
Ilja.kroeker@iws.uni-stuttgart.de
Assem Mayar
assem.mayar@iws.uni-stuttgart.de
Luca Trevisan
luca.trevisan@kit.edu
Holger Steeb
holger.steeb@mechbau.uni-stuttgart.de
Silke Wieprecht
Silke.Wieprecht@iws.uni-stuttgart.de
Wolfgang Nowak
Wolfgang.nowak@iws.uni-stuttgart.de

\textsuperscript{1} Institute for Modelling Hydraulic and Environmental Systems (IWS), University of Stuttgart, 70569 Stuttgart, Germany
\textsuperscript{2} Institute for Hydromechanics (IfH), Karlsruhe Institute of Technology, 76131 Karlsruhe, Germany
\textsuperscript{3} Institute of Applied Mechanics (CE), University of Stuttgart, 70569 Stuttgart, Germany