**Glossary**

**accretionary lapilli:** see lapilli

**active volcano:** various uses of the term ‘active’ are found in the volcanological literature. For the purposes of this volume, we consider a volcano to be ‘active’ if it has erupted in the past and is likely to erupt at some time in the future. Thus, the time scale of interest is geological and is not restricted to human memory or recorded history.

**Akaike Information Criterion (AIC):** for a parametric model with \( m \) parameters, the AIC is given by \(-2 \ln L_{m} + 2m\), where \( L_{m} \) is the maximized log-likelihood of the model. Ideally, models have large likelihood and few parameters, so a common model selection criterion is to adopt the model with largest AIC. One application of this procedure is to assist the identification of the order of a time-series model.

**aleatoric uncertainty:** uncertainty that is explicitly recognized in a statistical process. The variation in a random variable and uncertainty in the estimation of the expected value of a parameter are examples of aleatoric uncertainty. In hazard assessment, aleatoric uncertainty is sometimes referred to as irreducible. For example, the exact number of volcanic eruptions that will occur in a given time interval in the future is unknown, even if the statistical process is known to be Poissonian and the recurrence rate is estimated. Usually used in context with epistemic uncertainty.

**andesite:** magma or rock crystallized from magma that is characterized by intermediate silica concentration. Andesites have rheological properties that are intermediate between those of basalt and rhyolite magmas. Silica content in andesites ranges from c. 52 to 68 wt%. Common minerals in andesite include fine-grained plagioclase, and occasionally amphibole and/or pyroxene. Andesite is extremely common in subduction zones, where tectonic plates converge and volatile compounds are introduced into the mantle.

**aquifer:** a body of rock that is porous enough to provide a supply of groundwater.

**ARIMA\((p, d, q)\) process:** ARIMA means autoregressive integrated moving average and describes non-stationary time series as made up of the sum of three components: \( p \) an autoregressive process in which the current value of the process is expressed as a finite linear aggregate of the previous values and an error; \( d \) a differencing degree needed to transform a non-stationary time series to a stationary one; \( q \) a moving average process that can be expressed as a weighted sum of white noise error. To find the most appropriate ARIMA model it is necessary to examine the autocorrelation and the partial autocorrelation of the time series. An ARIMA\((p, d, q)\) process is a process whose \( d \)th difference is an ARIMA\((p, q)\) process.

**ARMA process:** an ARIMA\((p, q)\) process is a linear combination of an autoregressive process of order \( p \) and a moving average process of order \( q \).

**aseismic:** characterized by a lack of seismic activity, or lack of anomalous seismic activity.

**ash:** pyroclast (fragment of volcanic rock) < 2 mm in mean diameter. Also see tephra.

**autocovariance:** see autocovariance.

**autocorrelation:** see autocovariance.

**autocovariance:** for a stationary time series \( X_{1}, X_{2}, \ldots \) the autocovariance function is \( \gamma(\tau) = \text{cov}(X_{t}, X_{t+\tau}) \). The normalized function \( \rho(\tau) = \gamma(\tau)/\gamma(0) \) is termed the autocorrelation function. Both these functions measure the strength of linear dependence in a time series at different lags \( \tau \). One interpretation of the autocovariance function is as the Fourier transform of the spectrum.

**autoregressive process:** an autoregressive process of order \( p \) is a time series \( X_{1}, X_{2}, \ldots, \) in which \( X_{t} \) is a linear combination of \( X_{t-1}, X_{t-2}, \ldots, X_{t-p} \) plus an error term.

**autoregressive spectrum:** spectrum associated with an autoregressive model of order \( n, \text{AR}(n) \). Obtained by converting the model into a complex number form that is dependent on the frequency \( \omega \) by letting \( z^{-\tau} = \exp(-j\tau\omega), \ r = 1, 2, \ldots, n \) where \( j \) is the complex number. This yields the following estimate of the spectral density function:

\[
\hat{h}(\omega) = \frac{\sigma^{2}}{2\pi} \frac{1}{1 + \hat{a}_{1} \exp(-j\omega) + \cdots + \hat{a}_{n} \exp(-jn\omega)}
\]

where \( \hat{a}_{i}, i = 1, 2, \ldots, n \) are the AR\((n)\) parameter estimates. This expression can be used to compute the amplitude and phase for a user-selected range of frequencies \( \omega \).

**backward shift operator** \((z^{-1})\): if a variable \( y(t) \) is sampled uniformly at a sampling interval of \( \Delta t \) time units, then the value at the \( k \)th sampling instant is...
denoted by \( y_k = y(k \Delta t) \). The operator \( z^{-r} \) shifts this sampled variable backward by \( r \) sampling intervals; that is, \( z^{-r} y_k = y_{k-r} \), where \( r \) can take on any integer value.

**basalt**: a black volcanic rock containing not more than about 52% \( \text{SiO}_2 \) by weight. Basaltic lava has a low viscosity. Volcanic gases can escape easily without generating enormous Plinian eruption columns. Basalt is erupted at temperatures between 1100 and 1250°C. Basalt flows cover about 70% of the Earth’s surface and huge areas of the terrestrial planets, and so are possibly the most important of all crustal rocks.

**Bayesian statistics**: a paradigm for statistical inference that hinges on the specification of prior distributions for all unknown parameters, followed by an application of Bayes theorem to incorporate the extra information included in data. The use of Bayesian techniques has historically been clouded in controversy because of the subjective nature of prior specification. However, recent developments in computational techniques have led to so many advantages for the Bayesian approach that many opponents have been won over. Advocates of the Bayesian view of statistics would argue that the requirement to include additional information in the form of prior distributions is in any case an advantage of the paradigm.

**Bayes theorem**: for events \( A \) and \( B \) with \( \Pr(B \neq 0) \), this theorem states \( \Pr(A | B) = \Pr(B | A) \Pr(A)/\Pr(B) \). Its role is to allow the reversal of conditioning of events in probability statements. In the context of Bayesian statistics this theorem is often applied with \( A \) representing model parameters and \( B \) representing data.

**bias**: in the estimation of a parameter \( \theta \) with a statistic \( T(X_1, \ldots, X_n) \), the bias is given by \( E[T(X_1, \ldots, X_n)] - \theta \). In other words, the bias is the average deviation of an estimator from the true parameter value.

**Bingham liquid**: a liquid that does not strain in response to an applied stress until a critical yield stress is reached. Behaves like a Newtonian liquid if applied stress is higher than critical yield stress.

**block**: pyroclast (fragment of volcanic rock) >64 mm in mean diameter that is comparatively cool and that does not deform ductily in flight or on impact. Blocks often break on impact with the ground surface. Also see tephra and bomb.

**bomb**: pyroclast (fragment of volcanic rock) >64 mm in mean diameter that is sufficiently hot to deform ductily in flight or on impact. Also see tephra and block.

**bubbly flow**: a flow regime characterized by bubbles dispersed in a continuous fluid phase. In volcanology, bubbly flow is most often associated with magma flow in conduits at pressures low enough to result in supersaturation of volatiles, and at pressures high enough to prevent dispersed flow. Bubbly flow also occurs in lavas.

**bulk modulus**: in an elastic closure model, a coefficient relating volume pressure change.

**caldera**: a basin-shaped depression, often circular, formed at the summit of a shield volcano or composite volcano, or occurring in an area independent of these structures. Calderas form when magma is withdrawn or erupted from a shallow magma chamber in sufficient volume for the overlying rock to collapse, resulting in a depression. Calderas range in diameter from a few kilometres to >40 km. Domes, cinder cones and composite volcanoes often grow within calderas.

**causal signal**: in signal processing, a signal taking zero values for any negative time: \( s(t) = 0, t < 0 \). A signal taking zero values for any positive time is called anti-causal, and signals taking non-zero values in both positive and negative times are called non-causal.

**central limit theorem**: see normal distribution.

**cinder cone**: a small volcano, typically <1 km in diameter at its base, formed by eruption of pyroclastic material, such as scoria, ejected from a vent. Cinder cones are typically constructed from Strombolian and sub-Plinian eruptions, and are often surrounded by lava flows erupted from the same vent. Examples of cinder cones include Paricutin, Mexico, and Cerro Negro, Nicaragua.

**Clark–Evans test**: a statistical test used to determine if a spatial point process can be modelled as a completely spatially random process (i.e. a Poisson process), based on nearest-neighbour distances. For a Poisson process, the expected spatial intensity might be estimated as \( \lambda = N/R \), where \( N \) is the number of observations that fall within a region of area \( R \). The expected value of the distance to nearest-neighbour point is \( E(X) = 1/2\sqrt{\lambda} \), and the variance is \( \text{Var}(X) = (4 - \pi)/4\pi\lambda \). If the observed distance to nearest-neighbour point is \( \tilde{X} \), then the z-score, \( Z = [\tilde{X} - E(X)]/\sqrt{\text{Var}(X)} \) can be compared with a standard normal distribution, to test for complete spatial randomness.

**Classical Model (Delft Classical Model)**: the title given to the underlying model of a calibrated procedure for combining expert opinions, developed at the Technical University of Delft, Netherlands, in which the mathematical basis for ascribing distinctive, performance-based, weights to participating experts is similar in concept to a conventional statistical significance test (hence ‘Classical’ Model). See also scoring rule and proper scoring rule.

**closure model**: additional to conservation laws, relations between independent random variables in a physical system. Examples include equation of state, crystal growth law.
**coda:** the later part of a seismic signal, particularly if it has a decaying amplitude. In many cases, the amplitude decays exponentially.

**coefficient of determination:** in a simple linear regression model, this is the square of the correlation coefficient, and corresponds to the proportion of variance in the model explained by the regressor variable. In more general models it maintains its definition as the proportion of variation in a response variable accounted for by the explanatory variables.

**co-ignimbrite plume:** see co-pyroclastic-flow plume.

**composite volcano (syn. Stratovolcano):** see volcano.

**conduit:** a channel, often pipe-like, through which magma flows toward the Earth’s surface. Volcanic conduits are often approximately cylindrical and typically a few tens of metres across. Conduits can be vertical or inclined.

**confidence interval:** in the estimation of a parameter \( \theta \), a \( q\% \) confidence interval is any interval \( [\theta_l, \theta_u] \) obtained on the basis of observed data such that \( \Pr(\theta \in [\theta_l, \theta_u]) = q\% \). Care is needed in interpretation, as it is the interval that is random (being dependent on the data) and not \( \theta \), so a correct interpretation is that in repeated applications \( q\% \) of intervals so calculated will contain the parameter \( \theta \) and \( (1-q)\% \) will not. The choice of \( q \) represents a question of balance: high confidence implies wide intervals; small intervals have low confidence. Depending on the context, 95% or 99% confidence intervals are often adopted.

**contingency table:** a frequency table showing the counts of two or more cross-classified categorical variables. The patterns of frequencies are often studied to examine the evidence of dependence between the variables.

**continuous wavelet transform (CWT):** see wavelet transform.

**co-pyroclastic-flow plume:** any buoyant ash plume generated from ejection atop pyroclastic flows irrespective of the nature of the pyroclastic flow. Co-pyroclastic-flow plumes include ash plumes from both dome-collapse and fountain-collapse pyroclastic flows. In particular, co-pyroclastic-flow plumes also include co-ignimbrite plumes, which are specific plumes originating from very pumiceous pyroclastic flows that eventually form ignimbrite deposits.

**coregionalization:** geostatistics aims to describe or model regionalized variables. These numerical functions, space and/or time dependent, present a dual behaviour with erratic variations as well as more structured (i.e. correlated) variations linked to their genesis. This randomly structured behaviour is called a regionalization; and a coregionalization, in a multivariate framework, when cross correlation between regionalized variables is examined.

**corner frequency:** in general, upper and lower frequency that limit the frequency band of an instrument or filter within which the amplitude and phase of a signal is not affected between input and output. In seismology, the corner frequency of an earthquake is defined as the frequency where the constant value at low frequencies is intersected by an asymptote for higher frequencies showing a slope proportional to a negative power of frequencies (i.e. high frequencies are damped). The corner frequency is a parameter used to define the characteristics of an instrument or filter, and an earthquake, respectively.

**correlation coefficient:** a (normalized) measure of the linear association between two datasets or random variables. All correlation coefficients range from between \(-1 \) and \(1\) where \(-1/1\) means perfect negative/positive correlation, respectively, and a correlation of zero means no linear association. The correlation between two random variables is their covariance divided by each variable’s standard deviation. In terms of the expected values of two random variables \( X \) and \( Y \) the correlation coefficient is given by

\[
\rho(X, Y) = \frac{E(XY) - E(X)E(Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}}.
\]

For computational purposes, a simple way to calculate a sample correlation coefficient, an estimate of \( \rho(X, Y) \), is

\[
R = \frac{N \sum_{i=1}^{N} X_i Y_i - \left[ \sum_{i=1}^{N} X_i \right] \left[ \sum_{i=1}^{N} Y_i \right]}{\sqrt{N \sum_{i=1}^{N} X_i^2 - \left[ \sum_{i=1}^{N} X_i \right]^2} \sqrt{N \sum_{i=1}^{N} Y_i^2 - \left[ \sum_{i=1}^{N} Y_i \right]^2}}
\]

where \( N \) is the number of observations.

**covariance:** a measure of the linear association between two datasets or random variables. Given two random variables \( X \) and \( Y \) their covariance is given by

\[
\text{cov}(X, Y) = E((X - E(X))(Y - E(Y)))
\]

where \( E(\cdot) \) refers to the expected value of the random variable.

**cross covariance:** given two stationary time series, \( X_1, X_2, \ldots \) and \( Y_1, Y_2, \ldots \), the cross covariance function is \( \gamma(\tau) = \text{cov}(X_\tau, Y_{\tau+\tau}) \). It measures the strength of the dependence (not necessarily linear) between two time series as a function of the lag \( \tau \).

**cross variogram:** given two (not necessarily stationary) time series, \( X_1, X_2, \ldots \) and \( Y_1, Y_2, \ldots \),...
...the cross variogram function is
\[ V_{XY}(\tau) = \frac{1}{2} \text{cov}[(X_{t+\tau} - X_t), (Y_{t+\tau} - Y_t)]. \]

It measures the strength of the dissimilarity (not necessarily linear) between two time series as a function of the lag \( \tau \).

crust: the thin outermost solid layer of the Earth. It represents <1% of the Earth’s volume and varies in thickness from c. 5 km beneath the oceans to c. 60 km beneath mountain chains.

cryptodome: a shallow-level intrusion of magma that is close enough to the Earth’s surface to cause pronounced deformation, or doming, of the Earth’s surface.

crystallization: conversion of a silicate melt into crystals during solidification of magma.

cumulative distribution function, cdf: the cdf of a random variable \( X \) is given by \( F(x) = \Pr(X \leq x) \). For a continuous random variable \( X \) with probability density function \( f(x) \), this implies \( F(x) = \int_{-\infty}^{x} f(u)\,du \). The terms cumulative distribution function and distribution function are synonymous.

data clustering: in data analysis, a technique consisting of partitioning a data set into subsets (or clusters), so that the data in each subset (ideally) share some common feature, often similarity or proximity for some defined distance measure.

degassing n. (degus) v.: the process by which volatiles that are dissolved in the melt come out of solution in the form of bubbles. We can distinguish between open- and closed-system degassing. In the former, the walls of the system are permeable to volatiles, which can therefore be lost or gained by the system. In the latter, the walls are impermeable and so the total amount of volatiles in the bubbles and in solution in the magma is conserved.

delphi method: an early (non-Bayesian) model for combining expert opinions in long-range forecasting problems, developed by the RAND Corporation in the 1950s, which operated on experts’ point assessments of unknown quantities by iterating a specific, or best-fit, set of parameters. The downhill simplex algorithm uses a geometric construct, the simplex, to explore a specific, \( N \)-dimensional parameter space. The simplex is a geometric object, such as a triangle or tetrahedron, in which each vertex of the simplex is defined by a specific parameter set. In an \( N \)-dimensional solution space, the simplex has \( N + 1 \) vertices. The downhill simplex algorithm adjusts these vertices using some rule set to identify an optimal, or best-fit, set of parameters for a specific dataset.

dense rock equivalent (DRE): estimate of the original volume of erupted unvesiculated magma. Calculated as: \( V_{\text{magma}}(\text{DRE}) = \rho_{\text{dep}} \times V_{\text{dep}} / \rho_{\text{magma}} \), where \( V_{\text{dep}} \) is the volume of the deposit, \( \rho_{\text{dep}} \) is the density of the deposit and \( \rho_{\text{magma}} \) is the density of the magma.

derivative operator \( (D, s, p) \): the time derivative operator is defined as \( D^t = d^t/dt^r \); that is, \( D^t y(t) = d^t y(t)/d^t t \). As this is closely related to the Laplace operator \( (s, p) \), then these are often used to represent the derivative operator; that is, \( D^t y(t) = d^t y(t)/d^t t \).

diffusion, n. (diffuse, v.): movement of molecules, ions or particles from a region of higher to one of lower solute concentration as a result of their random thermal agitation.

diffusivity: a measure of the rate at which a particular species of volatile molecules moves down a concentration gradient in a particular material.

discursive Fourier transform (DFT): see Fourier transform.

discursive wavelet transform (DWT): see wavelet transform.

dispersed flow: a flow regime characterized by the presence of solid particles or liquid drops in a continuous gas phase. In volcano conduits, dispersed flow occurs about the fragmentation level, greatly accelerating flow velocity.

distribution function: see cumulative distribution function.

dome: a steep-sided, often bulbous extrusion (exogenous dome) or shallow intrusion (endogenous dome) of lava. Domes are frequently, but not exclusively, composed of magmas of more silicic composition. In dome-forming eruptions the erupted magma is so viscous, or the effusion rate so slow, that lava accumulates very close to the vent region, rather than flowing away. Pyroclastic flows can be generated by collapse of lava domes. Recent eruptions producing lava domes include the 1995–2001 eruption of the Soufrière Hills volcano, Montserrat, and the 2004–2005 eruption of Mount St. Helens. Also see cryptodome.

downhill simplex algorithm: an optimization, or nonlinear inversion, algorithm for estimating best-fit parameters. The downhill simplex algorithm uses a geometric construct, the simplex, to explore a specific, \( N \)-dimensional parameter space. The simplex is a geometric object, such as a triangle or tetrahedron, in which each vertex of the simplex is defined by a specific parameter set. In an \( N \)-dimensional solution space, the simplex has \( N + 1 \) vertices. The downhill simplex algorithm adjusts these vertices using some rule set to identify an optimal, or best-fit, set of parameters for a specific dataset.

dyadic DWT: see wavelet transform.

dyke: a sheet-like igneous intrusion that cuts across pre-existing, older, geological structure,
**GLOSSARY**

265

often vertical or near vertical. During magmatism, dykes transport magma toward the surface or laterally in fracture-like conduits. In the geological record, dykes are preserved as sheet-like bodies of igneous rocks.

**dynamic autoregression (DAR):** a dynamic autoregression model is a time-variable parameter autoregression model (see time-variable parameter estimation). This allows for the modelling of non-stationary stochastic time series $y_t$ with time-varying statistical characteristics and is useful for exploring the time–frequency characteristics, with the DAR defining the signal spectrum of $y_t$ at each sampling instant through time.

**dynamic harmonic regression (DHR):** a dynamic harmonic regression model is a harmonic regression model for a uniformly sampled non-stationary time series $y_t$, where the normally constant parameters defining the regression relationship are replaced by time-variable parameters (see also time-variable parameter estimation). This allows for the modelling of non-stationary periodic or quasi-periodic times series with time varying amplitude and phase characteristics.

**dynamic linear regression (DLR):** a dynamic linear regression model is a linear regression model where the normally constant parameters defining the regression relationship are replaced by time-variable parameters (see time-variable parameter estimation). This allows for the modelling of a regression relationship that changes in a statistically significant manner over time. (It should be noted that here and in other similar definitions within this Glossary, the term ‘dynamic’ is used because of historical precedence to mean a model with time-variable parameters. It is not the usual meaning of the word within a dynamic systems context.)

**dynamic transfer function estimation:** the estimation of the time-variable parameters in a non-stationary transfer function model (see time-variable parameter estimation). This allows for the modelling of non-stationary input–output behaviour and is useful for investigating systems whose dynamic characteristics change significantly over time.

**effusive eruption:** a non-explosive volcanic eruption in which coherent molten rock is extruded from the vent to form lava flows. Effusive eruptions do not involve fragmentation. Also see explosive eruption and extrusive flow.

**elastic:** in rock mechanics and seismology, the property of recovering original dimensions or volume after an applied stress is removed.

**elements at risk:** the population, buildings and civil engineering works, economic activities, public services, utilities and infrastructure, etc. at risk in a given area.

**elicitation facilitator:** a trained uncertainty analyst, essential for defining the issues and guiding the progress of an elicitation of expert opinion, and for processing the results.

**epistemic uncertainty:** uncertainty that arises from the choice of model or lack of knowledge about a physical system. For example, two different models of laharc run-out may rely on the same parameters, but use different physical assumptions to model the process. The resulting hazard forecasts may vary from one another, an example of epistemic uncertainty. In hazard assessment, epistemic uncertainty is often considered to be reducible by further model testing and improvement in our understanding of natural phenomena. Also see aleatoric uncertainty.

**eruption, n. (erupt, v.):** the process of magma reaching the surface and flowing from a volcano vent, either in an explosive manner, producing pyroclasts, or an effusive manner, producing lava flows. Eruptions exhibit considerable variation in magnitude, as reflected in the volcano explosivity index (VEI). See: Hawaiian-, Strombolian-, Vulcanian-, and Plinian-style eruption.

**eruption column:** cloud of pyroclastic material that rises above an erupting volcano. The eruption column is normally divided into a region of momentum-dominated flow immediately above the vent, and buoyancy driven flow in the upper part of the column. Eruption columns can reach elevations ranging from hundreds of metres to 40 km or more.

**estimate:** see estimator.

**estimator:** a statistic used in the estimation of a parameter or other quantity. The realized value of the estimator is termed the estimate. For example, to estimate a population mean $\mu$, the sample mean $(X_1 + X_2 + \cdots + X_n)/n$ is an (unbiased) estimator. Once the values of $X_1, \ldots, X_n$ are known, as $x_1, \ldots, x_n$ say, then $(x_1 + x_2 + \cdots + x_n)/n$ is the estimate of $\mu$.\n
**event:** statistically, one realization from the set of all possible realizations of a random variable. Physically, some specific natural observation, for example, an earthquake or a volcanic eruption may be an event.

**expected value:** the theoretical mean value of a random variable $X$. When $X$ is discrete with probability function $p(x)$, $E(X) = \sum xp(x)$. When $X$ is continuous with probability density function $f(x)$, $E(X) = \int xf(x)dx$.

**expert judgement:** the exercise of reasoning and decision-making within the framework of formal logic on the basis of specialist or domain-specific knowledge by an individual with acknowledged expertise in a subject; more specifically, the expert states his degree of belief in a proposition in terms of a ‘subjective probability’ for the truth or validity of that proposition, thereby giving quantitative expression to his uncertainty in relation to the matter in question.
**explosive eruption**: a volcanic eruption that involves fragmentation. These eruptions are powered by the rapid expansion of gas bubbles within the magma. The expansion is usually a combination of nucleation of bubbles and their subsequent growth by diffusion and decompression. It is also possible that some eruptions are powered primarily by decompression alone. The gas expansion is sufficiently rapid to cause the magma to be torn apart, that is, to be fragmented. The fragments are called pyroclasts.

**exponential distribution**: a random variable \( X \) has the exponential distribution with parameter \( \lambda \) if its probability density function is given by \( f(x) = \lambda e^{-\lambda x} \) for \( x > 0 \). One derivation of the exponential distribution is as the distribution of inter-arrival times in a homogeneous Poisson process with parameter \( \lambda \). It is therefore a canonical model for the distribution of times between randomly occurring events.

**exsolution, n. (exsolve, v.):** exsolution is the process by which volatiles come out of solution, that is, the process of bubble nucleation and growth.

**extreme value theory**: branch of statistics dedicated to the mathematical characterization, statistical modelling and inference of unusually large (or small) values of a stochastic process.

**extrusive flow**: a non-explosive (non-pyroclastic) magma flow from a volcanic conduit during a lava dome-building eruption or lava flow.

**fast Fourier transform (FFT)**: an efficient algorithm to compute the discrete Fourier transform and its inverse.

**fast wavelet transform (FWT)**: an efficient algorithm to compute the dyadic discrete wavelet transform and its inverse.

**fire fountain**: some lava eruptions are sufficiently violent that a momentum-controlled fluid fountain occurs above the vent. These fountains can feed lava flows.

**fissure**: a linear fracture on the Earth’s surface through which lavas, pyroclasts and gas may be erupted.

**Fourier transform**: an integral transform that expresses a function \( f(x) \) in terms of complex sinusoids:

\[ \mathcal{F}[f(x)] = \mathcal{F}(\omega) : \mathbb{C} \rightarrow \mathbb{C} = \int_{-\infty}^{\infty} f(x) e^{-i\omega x} \, dx \]

\( \omega \) is the angular frequency and is expressed in \( \text{rad s}^{-1} \). In signals processing, the transform is normally applied to a real signal \( s(t) \) and, instead of the angular frequency, the frequency,

\[ f = \omega / 2\pi, \]

expressed in Hz, is often preferred:

\[ S_F(f) : \mathbb{R} \rightarrow \mathbb{C} = \int_{-\infty}^{\infty} s(t) e^{-i2\pi ft} \, dt. \]

The Fourier transform of a causal sequence of finite length (such as, for example, a time series) is called the discrete Fourier transform (DFT):

\[ DFT \{ x[n] \} = X[k] : \mathbb{N} \rightarrow \mathbb{C} = \sum_{n=0}^{N-1} x[n] e^{-i(2\pi k/n)}, \quad 0 \leq k \leq N - 1. \]

**fragmentation, n. (fragment, v. or n.):** the process by which molten magma is disrupted into many fragments by internal degassing.

**gamma distribution**: a generalization of the exponential distribution, the gamma distribution can also be derived from a Poisson process. The time to the \( r \)th event in a Poisson process with parameter \( \lambda \) has the gamma distribution \( \text{Ga}(r, \lambda) \). The case \( r = 1 \) reduces naturally to the exponential distribution. The probability density function of a random variable \( X \) with distribution \( \text{Ga}(\gamma, \lambda) \) is

\[ f(x) = \frac{\lambda^\gamma x^{\gamma-1} e^{-\lambda x}}{\Gamma(\gamma)} \quad x > 0. \]

where \( \Gamma(\gamma) \) is the so-called gamma function, defined by

\[ \Gamma(\gamma) = \int_{0}^{\infty} u^{\gamma-1} e^{-u} \, du. \]

**Gaussian distribution** see normal distribution.

**generalized random walk**: general stochastic model used in time-variable parameter estimation. It contains, as special examples, the random walk (RW), the integrated random walk (IRW) and the smoothed random walk (SRW). In the simplest RW case, the model assumes that the parameter value at the \( k \)th sampling instant is equal to the value at the previous sampling instant plus a small, random change defined by a zero mean value, white noise input with specified variance.

**geological record**: see stratigraphic record.

**gliding spectral lines**: characteristic feature in a spectrogram of volcanic tremor observed at many explosive volcanoes. Spectral lines in a spectrogram are produced by the precisely repeated excitation of a source signal where the spectral distance between the lines is inversely related to the excitation frequency (number of events per time unit). Hence, gliding lines can be interpreted as a change in excitation frequency as observed before explosions (increasing number) or after episodes of degassing (decreasing number). Alternatively, the gliding of
lines can be due to changes in magma properties, producing changes in the seismic velocity. Superimposed on the gliding spectral lines are constant spectral features, which are due to the site characteristics of the seismic station.

**global minimum** the set of parameters that best minimize differences between model results and observations for a specific dataset. The minimum is global in the sense that no other parameter set within the solution space yields a better fit between model results and data.

**granular stress**: additional stress in a multiphase system as a result of particle interaction.

**Hawaiian-style eruption**: a type of volcanic eruption characterized by eruption of pyroclasts, often from fissure or vent systems that may extend for 1 km or more. Hawaiian eruptions from single vents typically have effusion rates of $10^2$ kg s$^{-1}$ or less, and pyroclasts reach <500 m above the vent. Hawaiian eruptions have a VEI of zero or one. Nevertheless, effusion rate and lava volume from Hawaiian eruptions can be large when integrated across an entire fissure zone and these eruptions can be sustained for a long time, commonly >1 year. Examples of Hawaiian-style eruptions are from Kilauea, USA, and the Laki fissure eruption, Iceland.

**hazard**: a physical or chemical process that has the potential to cause injury or loss of life to people or damage to property or infrastructure. See volcanic hazard.

**hazard function**: for a continuous random variable $T$ (generally with domain corresponding to time), $h(t)$ is the instantaneous failure rate at $t$ conditional on survival to $t$. More explicitly,

$$h(t) = \lim_{\Delta t \to 0^+} \frac{\Pr(t \leq T \leq t + \Delta t | T \geq t)}{\Delta t}.$$  

Because of the various relationships between the density, distribution and survivor functions, $h$ can also be expressed as $h(t) = f(t)/S(t)$, where $f$ and $S$ are the density and survivor functions, respectively.

**hidden Markov model (HMM)**: a model for a double stochastic process where the time-dependent phenomenon is associated with a Markov process, but the states over time are not directly observable. Instead, a second stochastic process, dependent on the states, is observable. In the simplest case, the set of values assumed by the observed state-dependent process is finite and the process can be described by a set of state-dependent probabilities. In the more general case, the observed state-dependent process can assume values in a continuous set of real numbers. It is then described by a set of probability density functions.

**Holocene**: a period of geological time defined as the interval from 10 ka before present to the present.

**homogeneous**: a stochastic process is homogeneous if its parameters remain constant over the domain. For example, in a homogeneous Poisson process the recurrence rate $\lambda$ is constant in time (or space).

**hybrid event (HY)**: a volcanic seismic event that shares some characteristics of both volcano-tectonic events and other event types, such as multi-phase events or long-period events. They may reflect a possible mixture of source mechanism or an intermediate location in the volcanic edifice. The typical duration is about 40 s.

**hyper-parameter**: generally used in a Bayesian setting to describe the parameters in the prior distribution of a model parameter.

**hypothesis test**: a statistical method to test some assumption about a parameter.

**igneous**: adjective used to describe characteristics pertaining to rocks that have crystallized from a magma.

**ignimbrite**: a deposit formed from pumiceous pyroclastic flows.

**independent random variable**: see random variable.

**instrumental variable estimation**: in linear regression analysis, it is assumed that the regression variables are exactly known. In its simplest form, instrumental variable (IV) estimation is a modification of least squares regression analysis in the situation where the regression variables are not known exactly but are affected by noise: the so-called ‘errors-in-variables’ problem. The IV is a variable that is correlated with the underlying noise-free values of the regression variables but uncorrelated with the noise that contaminates them. Having access to such a signal allows the ‘normal equations’ of regression analysis to be modified in a manner that removes the asymptotic bias that affects the least-squares estimates when errors in variables are present. It is particularly useful in the case of transfer function models, where the input variables provide a source of instrumental variables.

**integral wavelet transform (IWT)**: see wavelet transform.

**intensity function**: measures the rate of occurrence of points in a point process. If the process is homogeneous, the intensity function is constant. If the process is non-homogeneous, it will be a function of position, and will take larger values in regions where points are more likely. Technically, if $E(N(A))$ denotes the expected number of points of a point process $N(.)$ in a region $A$, then $E(N(A))$ is the integrated intensity function, of which the intensity function is the density.

**inversion**: a mathematical or algorithmic procedure for deducing optimal parameters that
relate some model to some dataset. In linear inversion, algebraic techniques are used to relate model parameters to data using a set of linear equations. In nonlinear inversion, algorithms must be used to search a parameter space efficiently to identify best-fit parameters.

**isovariogram** locations. multivariate case (co-kriging) for the integration minimum uncertainty (quantified by an error following properties: (1) linearity (in general); (2) unbiasedness (i.e. no systematic error); (3) minimum uncertainty (quantified by an error variance). This method can be generalized in the multivariate case (co-kriging) for the integration of various types of data measured at different locations.

**lag** The difference between two time points. For example, if \( t \) and \( s \) are two different time points with \( \tau = t - s \) then \( X_s \) is a lagged version of \( X_t \) with lag \( \tau \).

**lahar:** a flow of poorly sorted heterogeneous debris, primarily consisting of volcanic rocks of all sizes mixed with water, sometimes referred to as mudflows. Temperatures of lahars are usually <100 °C. Lahars are formed during volcanic eruptions; for instance, by eruption of material through a crater lake or melting of a volcano glacier. Lahars also occur following cessation of volcanic activity on unstable slopes, and are often triggered by rainfall.

**lapilli:** pyroclast (fragment of volcanic rock) >2 mm and <64 mm in mean diameter. Also see tephra. Lapilli are sometimes formed in eruption columns by the accretion of ash-sized particles, termed accretionary lapilli.

**Laurent transform:** see Z-transform.

**Lava:** molten rock erupted by a volcano, as an effusive flow at the Earth’s surface.

**lava dome:** see dome.

**leave-one-out method:** in pattern recognition, a method to estimate the error rate of a classification scheme. In this technique, the training of the classifier is performed using \((E - 1)\) events, \( E \) being the number of events at disposal. The test is then carried out using the excluded event. If this event is misclassified, an error is counted. The procedure is repeated \( E \) times, each time excluding a different event; in so doing all samples are used for both the training and the classification. This method gives an upper bound for the true error at the expense of a high computational cost.

**leucogranitic:** a silicic magma high in aluminium (peraluminous), often characterized by abundant biotite minerals in an otherwise pale groundmass.

**likelihood or likelihood function:** in statistics, a conditional probability function considered a function of its second argument with its first argument held fixed, and also any other function proportional to such a function:

\[
\mathbb{L}(b \mid A) : B \rightarrow \mathbb{R} = \alpha \mathbb{P}(A \mid B = b)
\]

for any constant of proportionality \( \alpha > 0 \). \( \mathbb{L}(b \mid A) \) is the likelihood of \( b \) relative to \( A \) and \( B \) is the set including the parameter \( b \). For instance, if \( A \) is an observed phenomenon and \( B \supseteq b \) is a set of models describing a phenomenon, \( \mathbb{L}(b \mid A) \) expresses how well the model \( b \) describes the observed phenomenon \( A \). It should be note that, unlike probability, the likelihood function can take any real value.

**local minimum:** a set of parameters that best minimize the differences between model results and observations within some local region of the parameter space. Occurrence of local minima
sometimes prevent nonlinear inversion algorithms from identifying global minima.

**logit**: a transformation where, if \( p \) is a probability,

\[
\text{logit}(p) = \log \left[ \frac{p}{1 - p} \right].
\]

**log-logistic distribution**: a standard distribution characterized by two parameters, the expected value, \( \mu \), and \( \kappa \). The log-logistic density, \( f_X(x) \), survivor, \( S_X(x) \), and hazard, \( h_X(x) \), functions for a random variable, \( X, x > 0 \), are

\[
f_X(x) = \frac{\kappa \mu^\kappa }{1 + (\mu x)^\kappa},
\]

\[
S_X(x) = \frac{1}{1 + (\mu x)^\kappa},
\]

\[
h_X(x) = \frac{\kappa \mu^\kappa }{1 + (\mu x)^\kappa}.
\]

It should be noted that if \( \kappa < 1 \) the hazard decreases with time. If \( \kappa > 1 \), the hazard goes through a single maximum, meaning that the hazard is greatest at some value of \( x \).

**log-normal distribution**: if \( X \) has the normal distribution \( N(\mu, \sigma^2) \) then the variable \( Y = e^X \) is said to have the lognormal distribution with parameters \( \mu \) and \( \sigma \). Having positive domain, the log-normal distribution is sometimes a more suitable model than the normal model for natural phenomena.

**long-period earthquake (LP)**: volcanic earthquake in a frequency range between 0.2 and 5 Hz, which is generated by a resonance of seismic energy that is trapped in a fluid-filled container embedded in a solid. Such earthquakes are composed of slow waves formed at the interface between fluid and solid, which are responsible for the low-frequency content. Models to generate these waves range from gas-filled cracks to magma-filled conduits. LPs are considered to be indicative of the pressurization and movement of magma.

**low-frequency earthquake (LF)**: earthquake category that includes long-period earthquakes together with so-called hybrid events, the latter having a high-frequency onset. On several volcanoes including Soufrière Hills on Montserrat, the same seismic event can be recorded at one station as a hybrid event and at another station as a long-period earthquake. As the amount of high-frequency energy added to the LP can vary over a wide range, both event types, hybrid and LP, are summarized as low-frequency earthquakes.

**mafic**: magma, lava, or tephra that is particularly rich in Fe - and Mg - bearing minerals, such as olivine and pyroxene. Often used to describe lava with silica concentrations of SiO\(_2\) < 55 wt%. Mafic magmas often have lower viscosity and lower volatile content than silicic magmas. Basalt is an example of a mafic magma.

**magma**: a hot melt containing dissolved volatiles and suspended crystals, which is generated by partial melting of the Earth’s crust or mantle and is the raw material of all volcanic processes. Silicate magmas are the most common magma type and consist of long, polymeric chains and rings of Si–O tetrahedra, between which are located randomly positioned cations (e.g. Ca\(^{2+}\), Mg\(^{2+}\), Fe\(^{2+}\) and Na\(^{+}\)) and anions (e.g. OH\(^-\), F\(^-\), Cl\(^-\) and S\(^-\)) that are loosely linked to the oxygens in the silicate tetrahedra. The greater the silica (SiO\(_2\)) content of the magma, the more chains and rings of silicate tetrahedra there are to impede each other and hence increase the viscosity of the magma. The pressure regime and composition of the magma control which minerals nucleate and crystallize from a magma when it cools.

**magma chamber**: a subsurface volume within which magma accumulates, differentiates and crystallizes. Granite bodies and similar structures illustrate the form of some magma chambers, but in general the shape and volume of magma chambers beneath active volcanoes are poorly known. Often magma reservoir is a better term.

**magmastic pressure**: the pressure related to the weight of a column of magma in an open volcanic conduit.

**Mallat’s algorithm**: a fast wavelet transform for computing dyadic DWTs introduced by S. G. Mallat in 1989.

**mantle**: zone lying between the Earth’s crust and core, c 2300 km thick. It represents about 84% of the Earth’s volume and 68% of its mass.

**Markov chain**: a discrete Markov process defined on a 1D domain (e.g. a time series).

**Markov model**: a special stochastic model, in the simplest case, a Markov model is a stochastic state model constrained by a simplifying condition called Markov’s property: the transition probabilities depend only on the current state and position in the domain. They do not depend on the history of the phenomenon. More strictly, if the transition probabilities depend only on the current state (and not on the position in the domain), the discrete Markov model is said to be homogeneous. For instance, if \( S_k \) is the state assumed at the time \( t_k \) by a homogeneous Markov chain, the transition probability from the state \( S_{k-1} \) to the state \( S_k \) after
having been in the states $S_{k-1}, S_{k-2}, \ldots, S_1$ is

$$p_{k-1, k} = P(S_{k-1} \rightarrow S_k \mid S_{k-1}, \ldots, S_1) = P(S_{k-1} \rightarrow S_k \mid S_{k-1})$$

whereas for a non-homogeneous Markov chain the transition probability depends also on the times $t_k$ at which the transitions occurred:

$$p_{k-1, k} = P(S_{k-1} \rightarrow S_k \mid S_{k-1}, t_{k-1}).$$

These models are Markov models of the first order. In the more general case, the transition probabilities depend on the previous $m$ transitions and the Markov model is said to be of order $m$. For a homogeneous Markov chain of order $m$ the transition probability from the state $S_{k-1}$ to the state $S_k$ is

$$p_{k-1, k} = P(S_{k-1} \rightarrow S_k \mid S_{k-1}, \ldots, S_{k-m})$$

whereas for a non-homogeneous Markov chain of the same order it depends also on the times $t_k$ at which the transitions occurred:

$$p_{k-1, k} = P(S_{k-1} \rightarrow S_k \mid S_{k-1}, t_{k-1}, \ldots, t_{k-m}).$$

**Markov process**: a stochastic process that can be described by a Markov model. In a time series $X_t, X_{t-1}, \ldots$ the distribution of $X_t$ is generally dependent on the entire history of the process. In a first-order Markov process the distribution of $X_t$ is independent of the history of the process, given the value of $X_{t-1}$. More explicitly, $P(X_t \mid X_{t-1}, X_{t-2}, \ldots, X_1) = P(X_t \mid X_{t-1})$. In a volcanic setting, $X_t$ might represent magma temperature on day $t$, and the Markov property would imply that knowing the temperature on day $(t-1)$, the temperature on previous days is irrelevant to the distribution on day $t$. More generally in a $p$th-order Markov process, $P(X_t \mid X_{t-1}, X_{t-2}, \ldots, X_1) = P(X_t \mid X_{t-1}, \ldots, X_{t-p})$. Markov processes often provide convenient models for dependence in time series, offering a balance in the need for models that provide sufficiently rich dependence structures, but that are simple enough to facilitate computation. This latter aspect is supported by a rich collection of theoretical results on Markov processes.

**Markov tree**: a Markov process defined on a tree-structured domain.

**maximum likelihood estimate**: a statistical method used to make inferences about parameters of the underlying probability distribution of a given dataset. The likelihood function $L(b \mid A)$ is maximized over all possible values of the parameter $b$. The value $b$ which maximises the likelihood function is known as the maximum likelihood estimator for $b$. In a parametric statistical model, the maximum likelihood estimate is the parameter estimate that attaches highest probability (or probability density) to the observed data.

**mean**: see sample mean and expected value.

**melt**: liquid part of magma. Magma may consist solely of melt, or of melt with suspended crystals and bubbles. Melts contain variable amounts of dissolved volatiles. The primary volatile is usually water.

**Message Passing Interface (MPI)**: a system of commands and utilities that enhance communication between different processors during parallel computation. MPI commands are included within parallel code, and are used to distribute information, such as data, problem domains or boundary conditions, among computer processors during execution of parallel programs and to gather information from processors.

**method of moments**: general procedure for estimating model parameters, in which the idea is to equate theoretical and sample moments of a model. For example, one equation is obtained by equating the theoretical mean of a model to the sample mean; a second by equating the theoretical variance to the sample variance. In a model with $p$ unknown parameters, equating the first $p$ theoretical and sample moments in this way usually generates $p$ equations, whose solution usually leads to an estimate of the $p$ parameters. The method is widely applicable, but has no general optimality properties.

**model**: a symbolic representation of reality. See also stochastic process and statistical model.

**monogenetic**: a volcano constructed from a single episode of volcanic activity, sometimes consisting of numerous eruptions, and that then never erupts again. There is no canonical distinction between monogenetic and polygenetic volcanoes established in the volcanological literature. Mono- genetic volcanoes are generally thought to be active for periods less than $c. 100$ years, and/or have no evidence of longer repose intervals between eruptions.

**Monte-Carlo methods**: generic term for procedures that allow the study of properties of statistical models via simulation. Commonly, realizations of a statistical model are generated on a computer, from which empirical estimates (e.g. sample mean) can be used to approximate theoretical equivalents (e.g. expected value).

**moving average process**: a time-series probability model where a process is constructed that is a finite linear combination of elements of a purely random process. The process cuts off at a lag equal to the order of the MA process.
multiphase event: a volcanic seismic signal presenting apparently many seismic phases. It is normally characterized by a low signal-to-noise ratio and it shows a dominant frequency content between 3 Hz and 4 Hz. It is associated with very shallow events and is probably related to dome growth. The shape could also be dependent on local scattering effects.

Newtonian liquid: a liquid for which the strain rate is proportional to the applied stress, where the proportionality coefficient is called the viscosity.

noise variance ratio (or signal-to-noise parameter): precise definition is model specific, but in general terms the noise variance ratio is the ratio of the variance of errors around a signal to the variance of the signal itself.

nonparametric: a branch of statistics using techniques that avoid the specification of parametric models. A common example is the kernel method for density estimation. Nonparametric methods are generally easier than parametric methods, but usually require heavier computation. They are also less efficient than parametric techniques when suitable parametric models can be found. Identification of relationships between variables is also generally easier through parametric models, but nonparametric methods are often useful as an exploratory first stage.

non-stationary: see stationary.

normal distribution: a random variable X is said to have a normal or Gaussian distribution with mean \( \mu \) and variance \( \sigma^2 \), denoted \( N(\mu, \sigma^2) \) if its probability density function is

\[
 f_X(x) = \frac{1}{\sigma \sqrt{2\pi}} \exp \left[ -\frac{(x - \mu)^2}{2\sigma^2} \right].
\]

The special case of \( \mu = 0, \sigma = 1 \) is referred to as the standard normal distribution. Computations on normal distributions are often simplified by applying a z-transformation: if X has distribution \( N(\mu, \sigma^2) \) then \( Z = (x - \mu)/\sigma \) has distribution \( N(0,1) \). The normal distribution has a special role in statistics as a result of the central limit theorem, which states that, under certain conditions, means or sums of sufficiently many independent and identically distributed random variables have a distribution that is approximately normal. For this reason the distributions of many physical phenomena might reasonably be modelled as normal. Countless studies in all fields of science lend empirical support to this theoretical argument. Nonetheless, the normal distribution is not a panacea for all statistical problems, and even when there is some theoretical justification for its use, empirical checks should be made to justify its use.

nucleation, n. (nucleus, n. sing.; nuclei, n. plu.): magmatic melts contain variable amounts of dissolved volatiles. When a melt becomes supersaturated, the volatiles will eventually come out of solution in the form of gas bubbles entrapped within the melt. According to classical nucleation theory, local fluctuations in the concentration of the dissolved volatile in the melt allow the formation of small clusters of molecules, the ‘nuclei’. To establish whether the nuclei disintegrate or lead to the formation of bubbles, we must consider the balance of the energies involved; the gas phase has a lower bulk free energy than the supersaturated melt, but energy is required to form the bubble interface. Above a certain size, the nuclei are stable; that is, the energy required to form the additional interface is less than the reduction in the bulk free energy and so the bubbles can grow spontaneously. The energy required for the formation of a critical nucleus (a nucleus of the size where the energies balance exactly) is called the ‘activation energy’. Bubble nucleation and specifically the distinction between ‘homogeneous’ and ‘heterogeneous’ nucleation is discussed in more detail in the first paper of this volume.

optimization: a set of mathematical procedures for estimating parameters that yield the best fit, or best performance, of a given model when applied to a specific dataset. Also see inversion.

parallel computation: the writing or use of computer programs that execute on multiple computer processors simultaneously. The goal of parallel computation is to speed problem solution by distributing the problem among several processors. This is often achieved because of the repetitiveness of many computational problems.

parameter, n. (parametric, adj.): unknown quantity in a mathematical model. In the context of parametric statistics a model usually includes unknown parameters whose estimation, using data and other information, completes the model specification. For example, in a simple linear regression model the usual parameters are the constants that determine the equation of the linear relationship between the two variables, and a variance that expresses variation in observations around the line. Once data are available, the method of least squares or maximum likelihood can be used to estimate these parameters.

partial autocorrelation: the partial autocorrelation at lag \( \tau \) measures the (extra) association between \( X_t \) and \( X_{t+\tau} \) (averaged over \( t \)) taking into account any variation explained by lower orders of lag.

pattern recognition: the scientific discipline whose aim is the classification of objects into a finite number of categories. The objects that need
to be classified are called the pattern and the measurement used for the classification, the feature. The categories into which patterns sharing similar features are grouped are called classes.

**Peclet number**: ratio of the time scales of diffusion (thermal or chemical) and the characteristic ascent time.

**Periodogram**: the periodogram is an estimate of the spectrum. The raw periodogram is highly variable and needs to be smoothed to be useful.

**Phase**: in seismology, a phase usually indicates a wavelet, which can be recognized and identified at several stations, or may be clearly attributed to a particular source.

**Phreatic**: relating to water. In volcanology, phreatic is often used to describe eruptions that result from heating of groundwater in the subsurface, resulting in vaporization of groundwater and subsequent explosion, and that do not result in the eruption of magma. Alternatively, phreatic often is used to describe water–magma interaction in the subsurface.

**Phreatomagmatic**: an explosive eruption characterized by interaction of magma and groundwater, producing pyroclasts, pyroclastic surges and related features.

**Pleistocene**: a period of geological time from 1.8 Ma ago to 10 ka ago.

**Plinian-style eruption**: an explosive eruption of pyroclastic ejecta forming an eruption column characterized by sustained flow, which may be up to 40 km or more high and disperse ejecta over an area of 500–5000 km². A Plinian eruption may produce thick tephra deposits, or the eruption column may collapse to generate pyroclastic flows.

**Pliocene**: a period of geological time from 5.4 to 1.8 Ma ago.

**Poisson distribution**: this is a standard distribution for discrete random variables. The probability function of a random variable whose distribution is Poisson with parameter \( \lambda \) is

\[
p(x) = \frac{e^{-\lambda} \lambda^x}{x!}, x = 0, 1, 2, \ldots
\]

The distribution is closely connected to the Poisson process. The number of events that occur in a unit interval for a Poisson process with parameter \( \lambda \) has a Poisson distribution with parameter \( \lambda \). Because the Poisson process is a reasonable stochastic description for many physical phenomena, this implies that the Poisson distribution is a plausible model for the distribution of counts of many types of physical events.

**Poisson process**: canonical stochastic model for the realization of point events in space or time (or both). The defining characteristic of a Poisson process is that point occurrence neither inhibits nor encourages occurrence of other points. In this sense, it is the standard model for complete randomness of points in space (or time).

**Polygenetic**: a volcano that is constructed from numerous episodes of volcanic eruption, some of which follow prolonged periods of repose, is termed polygenetic. Polygenetic volcanoes typically have total durations of volcanic activity of the order of \( 10^3 – 10^6 \) years. Also see monogenetic.

**Population mean**: see expected value.

**Probability density function (pdf)**: a measure of probability for all possible values of a random variable, or for any subset of them. For a continuous random variable, \( X \), a function, \( f_X(x) \), with the properties, \( \int f_X(x)dx = 1 \), and \( f_X(x) > 0 \) for all \( x \). Often the subscript \( X \), is omitted, as in \( f(x) \), when the random variable is clear from the context. Also see standard distribution.

**Proper scoring rule**: a scoring rule is termed proper if it has the property that, in order to optimize his score, it encourages the expert to express his true view, opinion, belief or subjective probability, otherwise he incurs a reduction in score (see scoring rule and classical model).

**Pumice, n. (pumiceous, adj.)**: an extremely vesicular (typically 60–80% volume of voids) rock that is formed in explosive eruptions. Pumices are often formed in magmas with very high silica (\( \text{SiO}_2 \)) contents such as rhyolites.

**Purely random process**: a stochastic process where the different random variables are independent and identically distributed. In mathematics if \( \{X_i\}_{i=1}^{\infty} \) is a purely random process then every \( X_i \) has the same probability distribution and all the \( X_i \) are statistically independent. Often used as a building block for other random processes.

**Pyroclast**: literally, ‘fire-broken’ (the prefix ‘pyro’ is from the Greek ‘pur’ meaning ‘fire’). Applied to volcanic rocks consisting of fragmented particles, generally produced in explosive eruptions.

**Pyroclastic flow or surge**: a gas-particle flow of pyroclasts suspended in hot air (or a mixture of hot air and ash). The flow originates by the gravitational collapse of a dense, turbulent explosive eruption column at the source vent, or by dome collapse, and moves down-slope as a coherent flow. Pyroclastic flows and surges are distinguished by particle concentration in the flow, surges being more dilute. This change in particle concentration results in differences in the deposits left by flows and surges.

**Quality factor, Q**: the quality factor, or damping factor, is a measure of the change in amplitude or energy of a signal over time at a single location, or in space at a single time or for a single phase. When used to describe the signal amplitude as a
function of time at a given location, it may be measured in the time domain by comparing the amplitude of one cycle of the signal with that of the next, or in the frequency domain by measuring the bandwidth of the spectral peak corresponding to the signal. The decrease in amplitude may be due to scattering or attenuation in the medium, or it may be due to damping (i.e., resistance to oscillation) in the source. The decay of a signal as it passes through space is usually measured in the time domain and is generally taken to be due to scattering and/or losses owing to friction.

**Quaternary**: a period of geological time defined as the interval from 1.8 Ma ago to the present.

**random variable**: a function that relates each result of an experiment to a real number. The random behaviour of some physical observations is quantitatively studied with a random variable. Two random variables, say \( X \) and \( Y \), are said to be **independent** if the conditional distribution of \( Y \) having observed \( X = x \) is identical to the marginal distribution of \( Y \). In other words, the variables are independent if the distribution of either is unaffected by knowledge of the value of the other.

**rarefaction wave**: propagating or stationary wave in which pressure decreases.

**recursive fixed interval smoothing**: in a recursive filtering algorithm (see **recursive state estimation**) the estimate at the \( k \)th sampling instant is based only on the data up to this instant and is usually denoted by \( \hat{x}_{k-1} \). On the other hand, the recursive fixed interval smoothing (FIS) estimate \( \hat{x}_{k|N} \) is based on the whole observational dataset of \( N \) samples. The FIS estimate is most easily generated by a backward recursive pass through the data that follows directly after the forward filtering pass. A similar FIS algorithm is used for estimating time-variable parameters. See **time-variable parameter estimation**.

**recursive parameter estimation**: in recursive parameter estimation, the parameters of a model are updated sequentially over time on the basis of the most recent data. In a typical recursive algorithm, the estimate \( \hat{a}_k \) of the model parameter vector \( a \) at the \( k \)th sampling instant is updated on the basis of the previous estimate \( \hat{a}_{k-1} \) and the ‘innovation’ or ‘one-step-ahead prediction’ error between the latest observation of the variable being modelled \( y_k \) and its predicted value \( \hat{y}_{k|k-1} \) based on the model incorporating the parameter estimate \( \hat{a}_{k-1} \).

**recursive state estimation**: in recursive state estimation, the state variables of a stochastic state-space model with known parameters are updated sequentially over time on the basis of the most recent data. In a typical recursive algorithm, the estimate \( \hat{x}_k \) of the \( n \)-dimensional model state vector \( x_k \) at the \( k \)th sampling instant is updated on the basis of the previous estimate \( \hat{x}_{k-1} \) and the ‘innovation’ or ‘one-step-ahead prediction’ error between the latest \( p \)-dimensional vector of observed variables \( y_k \) and its predicted value \( \hat{y}_{k|k-1} \) based on the predicted state estimate \( \hat{x}_{k|k-1} \). Here, \( \hat{x}_{k|k-1} \) is computed from the previous state estimate \( \hat{x}_{k-1} \) using the stochastic state–space model; and the observation \( y_k \) is a known linear combination of the state variables; that is \( y_k = Cx_k \), where the \( p \times n \) matrix \( C \) defines the nature of the linear combination. Recursive state estimation is also called ‘filtering’ and the most famous algorithm of this type is the ‘Kalman filter’ proposed by Rudolf Kalman in 1960.

**repose interval**: the time elapsed between successive volcanic eruptions at the same volcano. Ideally, repose interval would be the time elapsed from the end of one volcanic eruption to the beginning of the next. Unfortunately, eruption duration is rarely well documented. Therefore repose interval is often the time elapsed from the beginning of one eruption to the beginning of the next.

**Reynolds number**: the ratio of inertial forces to viscous forces acting on a fluid; used to determine if a flow will be laminar or turbulent. The Reynolds number is defined as

\[
Re = \frac{\rho u L}{\mu}
\]

where \( \rho \) is the fluid density, \( u \) is mean velocity, \( L \) is the characteristic length scale and is equivalent to the diameter for a circular conduit, and \( \mu \) is fluid viscosity. In particle settling, the Reynolds number can also be used to describe the flow of a particle through a fluid.

**reticulate**: an unusual pumice made of basalt (most pumices are formed in magmas with higher SiO₂ contents than basalt). An extremely vesicular (up to 95% volume of voids) rock that is produced during some Hawaiian-style eruptions.

**rheology**: the study of the flow properties of fluids.

**rhyolite**: a light-coloured volcanic rock with >68 wt% SiO₂. Sodium and potassium oxides both can reach about 5 wt%. Common mineral types include quartz, feldspar and biotite, and are often found in a glassy matrix. Rhyolite is erupted at temperatures of 700–850 °C.

**risk**: the probability of loss of life or damage to property or infrastructure.

**root-mean-square error (rms)**: the mean square error of an estimator, \( \hat{\mu} \), of a parameter \( \mu \) is

\[
MSE(\hat{\mu}) = E(\hat{\mu} - \mu)^2,
\]

and the root-mean-square error is the square root of \( MSE \). For a random variable \( X \), the variance of the estimated mean value of \( X \) gives the mean square error.
**Runge–Kutta method**: computational method for approximating the solution to ordinary differential equations, with initial conditions specified.

**sample**: in statistics, a collection of measures or results from experiments or observations. Often in geology, a piece of material, such as rock, about which observations are made or from which observations are drawn.

**sample mean**: the average of the sample measures. If the sample has not been drawn yet, a random variable. When the sample has been drawn, the sample mean value is an estimate of the expected value, a parameter. The sample mean of a random variable, $X$, is calculated as $\bar{X} = \frac{1}{N} \sum_{i=1}^{N} X_i$, where $N$ is the number of observations.

**scoria**: dark-coloured, vesicular pyroclastic fragments of basalt or andesite composition.

**scoring rule**: a formula for awarding a score (reward) to an expert in an elicitation, which is a function both of the expert’s elicited probability distribution for an uncertain quantity and of the true value of that quantity.

**seismogenic**: capable of, or having produced, earthquakes. In seismic hazard assessment, reference is often made to seismogenic zones, such as areas near faults or volcanoes, characterized by past seismic activity.

**settling velocity**: the terminal velocity of a particle falling through a fluid. For particles of tephra settling velocity depends on particle size (diameter), density and shape, and the density of the atmosphere.

**short-time average to long-time-average ratio (STA/LTA) trigger algorithm**: in seismology, a trigger algorithm that continuously calculates the average values of the absolute amplitude of a seismic signal in two consecutive moving-time windows. The short time window (STA) is sensitive to seismic events whereas the long time window (LTA) provides information about the temporal amplitude of seismic noise at the site. When the ratio of both exceeds a pre-set value, an event is declared.

**short-time Fourier transform (STFT):** a time–frequency transform using static time–frequency atoms that are windowed complex sinusoids:

$$S(t,f) : \mathbb{R}^2 \rightarrow \mathbb{C} = \int_{-\infty}^{\infty} s(\tau) w(\tau - t) e^{2\pi i f \tau} d\tau.$$

If the window $w(t)$ is a Gaussian function, the STFT coincides with the Gabor transform.

**silicic**: magma, lava, or tephra that is particularly rich in Al- and K-bearing minerals and silica concentrations of approximately SiO$_2 > 55$ wt%.

Silicic magmas are often more viscous and can have higher volatile contents than mafic magmas. Rhyolite is an example of a silicic magma.

**sill**: a sheet-like igneous intrusion that is concordant with pre-existing geological structure, often horizontal or nearly horizontal (see dyke).

**spatial process**: see stochastic process.

**spectrum**: the spectrum of a stationary time series is a function that describes how much variation is present in the time series at different frequencies. Alternatively, the spectrum is the Fourier transform of the autocovariance.

**standard distribution**: a probability density function that has common use in statistics. Such distributions often have a physical basis that makes them particularly appropriate for modelling statistical processes. See the following standard distributions described in this Glossary: Poisson, normal, lognormal, exponential, Weibull, gamma, log-logistic.

**state diagram or state transition diagram**: a diagram describing a finite state model (such as a stochastic state model), consisting of circles to represent states and directed line segments to represent transitions between the states. One or more actions (outputs) may be associated with each transition.

**stationary**: a process whose stochastic properties (mean, variance, marginal distribution, autocorrelation, etc.) remain constant over its domain is described as stationary. A process that is not stationary is termed nonstationary.

**statistical model**: mathematical model that includes some component of random variation. Typically, a statistical model includes a deterministic part, up to some unknown parameters, and a random part (possibly also with unknown parameters) that describes variation in observations around the deterministic part of the model. A simple example is a linear regression model, which comprises a linear component with unknown parameters plus random variation around the line. With observed data the usual objective is to estimate the deterministic component of the model, taking account of the random component.

**stochastic process**: an indexed set of random variables. A time series is a stochastic process whose index is (usually) the positive integers. A spatial process is a stochastic process whose index space maps to the region of interest.

**stochastic state model**: a model of a time-varying phenomenon obtained by associating it with a finite set of situations (or states) and then observing how it commutes from one state to another over the time. From this observation it is possible to estimate the state distributions (i.e. the probabilities of being in a given state) and the transition probabilities...
(i.e. the probabilities of changing from one state to another). These parameters are in general dependent on the time and on the states assumed by the phenomenon in the past. It should be noted that the current definition introduces the stochastic state model as a way to describe time-dependent processes. This is a specific case: in general, it can be defined on every kind of domain.

**strain**: change in shape, length or volume of material, such as rock, in response to stress.

**stratigraphic record**: stratification of sedimentary and/or volcanic deposits. The sequence of layers in the stratigraphic record gives an indication of relative time and duration of any given geological process. One of stratigraphy’s basic concepts is described by the Law of Superposition, which states that, in an undeformed stratigraphic record, the oldest layers (i.e. strata) occur at the base of the sequence. The law was first proposed in the 17th century by the Danish scientist Nicolas Steno.

**stratovolcano**: see composite volcano.

**stress**: force applied on a unit area of some piece of material, such as rock.

**Strombolian-style eruption**: a type of volcanic eruption characterized by discrete, often rhythmic, eruption of pyroclasts, driven by the release of gas at the very top of the volcano conduit. Gas is often released in coalesced slugs, sometimes >1 m in diameter. Eruption columns, characterized by the buoyant rise of elutriating tephra, rather than a sustained gas jet, rarely exceed 5 km and total tephra volume is usually <10^7 m^3 (mass of tephra ejected <10^11 kg). Effusion of lava flows often accompanies Strombolian eruptions. Examples of Strombolian eruptions are from Stromboli, Italy, and Izalco, El Salvador.

**supersaturation**: when the partial pressure of a dissolved volatile in magma exceeds the ambient pressure, the magma is supersaturated. A critical supersaturation has to be reached for bubbles to nucleate.

**support of a function**: the smallest closed subset of the domain outside of which the function and its first derivative are zero.

**survivor function**: complement of the distribution function, so if a random variable X has distribution function \( F(x) \), its survivor function is \( S(x) = \Pr(X > x) = 1 - F(x) \). The name is descriptive: \( S(x) \) is the probability of survival beyond \( x \).

Also see cumulative distribution function and hazard function.

**tectonic event**: an earthquake generated in response to plate tectonically induced stress.

**tephra**: all particles ejected from volcanoes irrespective of size, shape and composition. The term ‘tephra’ was introduced by Thorarinsson in the 1940s to describe particles that were projected through the air and eventually deposited on the ground, and therefore it includes the materials resulting from events described as tephra fall, air fall and pyroclastic fall, but not deposits resulting from flowage events. A variety of terms are used to describe the range of rock fragments erupted into the air by volcanoes. The terms classify the fragments according to size, shape, composition, or the way in which they form and travel.

**tidal potential**: for all planets in the Universe moving in a stationary orbit, the gravitational accelerations produced by other bodies (planets, satellites) are completely compensated in their centres of mass by the centrifugal accelerations owing to the orbital motion of the body. Any point on or within such a body, other than the centre of mass, experiences the difference between centrifugal and gravitational acceleration: the tidal acceleration. This difference is usually expressed as a potential, the partial derivatives of which are the tidal acceleration (vertical), tidal tilt and strain.

**time–frequency transform**: a transform that gives the frequency content of a signal as a function of time. The signal is decomposed in waveforms that are localized in time and frequency:

\[
S(t, f)|R^2 \rightarrow C = \int_{-\infty}^{\infty} s(\tau)w(\tau - t, f)d\tau
\]

where the real function \( s(t) \) is the signal to be analysed and \( \tau \) is the integration variable. The waveforms \( w(t, f) \) are called time–frequency atoms and are in general complex functions. The domain of the transformation is called the time–frequency plane.

**time series**: in statistics and signal processing, a set of data observations collected through time. For example, \( \{x_t\} = \{x_1, x_2, \ldots, x_{12}\} \) might be the monthly average temperature at a specific place in a given year (so \( x_1 \) might be the monthly average temperature for January, \( x_2 \) for February, and so on, until \( x_{12} \) for December). Time-series data are different from ‘elementary’ data as they are almost always correlated and have a definite order in time (e.g. in the above timeseries \( x_2 \) comes ‘later’ than \( x_1 \), and so on). A time series is not to be confused with a mathematical series, this being the sum of a sequence of terms.

**time-variable parameter estimation**: class of techniques for estimating statistical models whose parameters are time-dependent, for example, a regression model with coefficients that vary across the year.

**tornillo**: a particular class of transient volcanic seismic signal with a sinusoidal waveform and screw-like envelope profile that can last for several minutes. These signals are wavelets and have spectra with several very narrow peaks.
transfer function: complex function describing the linear characteristics of an instrument or filter in terms of frequency, where the complex characteristics are usually given in amplitude and phase. An instrument or filter is usually band-limited; hence, it does not change the amplitude or phase from input to output within a certain frequency band, given by the two corner frequencies. Outside this frequency band both the amplitude and the phase will be distorted, resulting in a change in waveform between input and output.

trellis diagram: an extension of the state transition diagram where the evolution of a finite state model over time is plotted. The circles represent the states which are repeated at every time unit for the detection of typical seismic signals (earthquakes, controlled source seismic signals, underground nuclear explosion signals, etc.) in the constantly present seismic noise signals.

variance: for a random variable \( X \) with mean \( \mu \), the variance is given by \( \text{Var}(X) = E(X - \mu)^2 \). As such, the variance measures the variability in a random variable. Consequently, if a random variable has small variance, repeat realizations will be similar; with a large variance they may be very different. Given data \( x_1, \ldots, x_n \), the variance is usually estimated by the sample variance

\[
s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2
\]

where \( \bar{x} \) is the sample mean. The factor \( (n - 1) \) in the denominator is a small correction to the more natural value of \( n \) and ensures that the sample variance is an unbiased estimator of the true population variance.

variogram: an alternative to the autocovariance function, the variogram measures the dissimilarity with lag of a time series. For a series \( X_1, X_2, \ldots \), the variogram is defined as

\[
V(\tau) = \frac{1}{2} \text{Var}[X_{t+\tau} - X_t].
\]

Its advantages over the autocovariance include the fact that it may be used for some nonstationary series, which is more natural for series whose time points are irregularly spaced, and that it is a natural analogue for spatial processes.

vent: the crater formed at the Earth’s surface by material emitting from, or collapsing into, the volcanic conduit.

very-long-period earthquake (VLP): sometimes also referred to as ultra-long-period, a type of volcanic earthquake with periods above 30 s. These earthquakes are in the transitional frequency band to ground motions with even longer time constants, usually referred to as ground deformation. Because of its (frequency-) band-limited property, a seismometer will alter the waveform of a VLP such that the recorded signal cannot be directly interpreted as ground motion. The extent of amplitude and phase distortion introduced by the instrument is dependent on its band-pass characteristics (frequency range) described by its transfer function.

vesicle, n. (vesicular, adj.): a bubble-shaped cavity in lava, pumice or scoria formed by the expansion of entrapped gases. A vesicular rock has many such cavities.

viscosity, n. (viscous, adj.): a measure of the resistance to flow of a material. Viscosity is usually defined as the ratio of the imposed stress to the rate of deformation (or ‘strain-rate’, which is the same as the spatial velocity gradient in the flow).

volatile: a dissolved component in a magmatic melt that would be gaseous at that temperature except for the confining pressure and solvent nature of the magma. The most common volatile in magmas is water vapour \( \text{H}_2\text{O} \), but there are often also significant quantities of \( \text{CO}_2 \) and \( \text{SO}_2 \).

Volcanic Explosivity Index (VEI): a classification scheme for volcano eruption magnitude, primarily defined in terms of the eruption column height and total volume of erupted material. VEI varies from zero (very small eruption) to eight (largest volcanic eruptions identified in the geological record).

volcanic hazard: the probability of occurrence, within a specific period of time in a given area, of a potentially damaging volcanic event.

volcanic risk: the expected number of lives lost, persons injured, damage to property and disruption of economic activity as a result of a particular volcanic event; consequently, volcanic risk is the product of volcanic hazard, vulnerability and elements at risk.

volcanic seismic signals: many of the processes occurring in active volcanoes generate vibrations that we can measure and record in a time series using seismometers and digitizers. Some of the signals are transients, that is short duration signals, whereas tremor continues for long intervals. Both these classes have many different appearances in the seismic record. Depending on their signal content, transients may be volcano-tectonic...
events, long-period or low-frequency earthquakes or tornillos. Tremor is also classified by its appearance as spasmodic, banded or harmonic.

volcanic tremor: superposition of low-frequency earthquakes that are repeatedly generated, leading to tremor episodes of durations between tens of seconds and days, and stable spectral characteristics. More generally referred to as seismic background noise on a volcano.

volcano: a naturally occurring vent or fissure at the Earth’s surface through which erupt molten, solid and gaseous materials. We can distinguish between ‘strato-’ or ‘composite volcanoes’ and ‘shield volcanoes’. A composite volcano is a large volcano, typically >1 km in diameter, principally formed from eruption of tephra and lava flows from a central vent. Composite volcanoes are characterized by episodes of volcanic eruptions, often separated in time by repose intervals that may last more than 100 years. Examples of composite volcanoes include Mt. Vesuvius, Italy, and Mount St. Helens, USA. By contrast, a shield volcano is often produced from Hawaiian-style eruptions that tend to produce a broad, low-angle cone, that is, Kilauea volcano, Hawaii, USA. Also see caldera, cinder cone, dome, monogenetic, polygenetic.

volcano-tectonic event: an earthquake that occurs in close proximity to a volcano and that can usually be attributed to stresses related to volcanic activity, but cannot necessarily be ascribed directly to the movement of magma in the subsurface.

Vulcanian-style eruption: a type of volcanic eruption characterized by discrete explosions, which produce shock waves, typically generated when exsolution of volcanic gas causes the shallow volcano conduit or dome to pressurize to the point of brittle failure. Volcanoes with recent Vulcanian eruptions include the Soufrière Hills volcano, Montserrat, and Colima volcano, Mexico.

vulnerability: the degree of loss to a given element at risk or set of such elements resulting from the occurrence of a given volcanic event, expressed on a scale from zero (no damage) to one (total loss).

wavelet: a single-sided signal of finite energy. That is, if we describe a wavelet as $f(t)$, with $-\infty < t < \infty, f(t) = 0$ for $t < 0$ (with $t = 0$ the start time of the wavelet), and with $\int_{-\infty}^{\infty}f(t)^2 dt < \infty$. In wavelet analysis, a general signal is decomposed into separate elements, in much the same way as sine or cosine waves are used in Fourier analysis, using a suitable, usually relatively simple wavelet, such as the Haar wavelet. In mathematics and signal processing, a wavelet refers to an oscillating function whose supports, both in the time and frequency domain, are closed and bounded. As a consequence, the energy associated with it is finite. In seismic studies, a wavelet is a 1D pulse, often the basic response of a single reflector.

wavelet transform: a family of time–frequency transforms using dynamic time–frequency atoms (called wavelets), which are translations and dilation of the same basic waveform $\psi(t)$, called the mother wavelet. The more general form, called the integral wavelet transform (IWT), is

$$S_{\psi}(b, a) : \mathbb{R}^2 \rightarrow \mathbb{C} = \int_{-\infty}^{+\infty} s(\tau) \frac{1}{\sqrt{a}} \psi^*(\frac{\tau - b}{a}) d\tau$$

where the operator $^*$ denotes complex conjugation. The parameter $a \in \mathbb{R} \setminus \{0\}$ is called the scale and the parameter $b \in \mathbb{R}$ is called the location. It should be noted that the IWT is actually defined on a time-scale plane: the location $b$ represents a translation in time of $\psi(t)$ and is measured in seconds, but the scale $a$ represents a dilation of $\psi(t)$ and is dimensionless. However, it can be shown that $S_{\psi}(b, a)$ gives local information about the frequency contents of the signal $s(t)$ in an ideal rectangle in the time–frequency plane centred on the point $(b/a, f_0/\omega)$ and whose extension on the plane is

$$\Delta t \times \Delta f = a \cdot L_0(f) \times \frac{L_M(f)}{a} = \text{const.}$$

$l_0$ and $l_M(f)$ are the centre and the length of the support of the mother wavelet $\psi(t)$, and $f_0$ and $L_M(f)$ are the centre and the length of the support of its Fourier transform $\Psi(f)$. This ideal rectangle is called a time–frequency cell.

If the scale and location parameters are discrete, so that the mother wavelet is scaled and translated only in discrete steps, the IWT becomes a discrete wavelet transform (DWT):

$$S_{\psi}(n, m) : \mathbb{Z}^2 \rightarrow \mathbb{C} = \int_{-\infty}^{+\infty} s(\tau) \frac{1}{\sqrt{m}} \psi^*(\frac{\tau - n}{m}) d\tau$$

where $m \in M \subseteq \mathbb{Z} \setminus \{0\}, n \in N \subseteq \mathbb{Z}$. It should be noted that the attribute ‘discrete’ refers only to the dilation $m$ and the translation $n$; the wavelet $\psi$ is in general a continuous function.

If the parameters $m$ and $n$ are chosen so that the length of the wavelet at scale $j + 1$ is twice its length at scale $j$, incrementing the scale halves the resolution in time and doubles the resolution.
in frequency. A DWT with the parameter chosen in this way is also called a dyadic DWT. In a dyadic DWT, the redundancy in frequency is minimized. In opposition to the DWT, the IWT is also often called a continuous wavelet transform (CWT).

Weibull distribution: a standard distribution that generalizes the exponential distribution, and is characterized by a parameter, \( \lambda \), where the expected value is \( 1/\lambda \), and an index parameter, \( \kappa \). This distribution was invented to describe material failure and has been used extensively in volcanology to forecast eruptions in terms of rock-mechanical failure. The Weibull density function \( f_X(x) \), survivor function \( S_X(x) \) and hazard function \( h_X(x) \) for a random variable, \( X \), \( x > 0 \), are

\[
\begin{align*}
  f_X(x) &= \kappa \lambda (\lambda x)^{\kappa-1} \exp[-(\lambda x)^\kappa] \\
  S_X(x) &= \exp[-(\lambda x)^\kappa] \\
  h_X(x) &= \kappa \lambda (\lambda x)^{\kappa-1}.
\end{align*}
\]

It should be noted that if \( \kappa = 1 \), the distribution is exponential. If \( \kappa > 1 \), \( h_X(x) \) increases with \( x \). Because of this property, the Weibull distribution is sometimes referred to as an accelerated failure model.

Z-transform: in mathematics and signal processing, a placeholder name for the Laurent transform which is a transform that converts a time series \( s[t] \) into a complex frequency domain representation:

\[
Z\{s[t]\} = S_Z(z) : \mathbb{R} \rightarrow C = \sum_{t=-\infty}^{\infty} s[t]z^{-t}
\]

where \( z = (a + ib) \) is a generic complex number. The unilateral Z-transform is to time series what the one-sided Laplace transform is to continuous time domain signals. The Z-transform computed on the set \( \{z = \exp(i\frac{2\pi k}{N})\} \), with \( k, N \in \mathbb{N} \) and \( 0 \leq k \leq N - 1 \), coincides with the discrete Fourier transform of a time series.

z-transformation: see normal distribution.
Index

Page numbers in italic, e.g. 79, refer to figures. Page numbers in bold, e.g. 123, signify entries in tables.

accretionary lapilli (see lapilli)
adaptive Gaussian kernel 80
defining events 80–81
modelling spatial intensity 81–84
advection–diffusion models 232–233
Akaiake Information Criterion (AIC) 148
aleatoric uncertainty 32, 36, 89, 240, 255
andesite 60, 96, 175, 181–182
Aragats, Mount 79, 80
Ararat, Mount 79, 80
Armenia, spatial distribution of volcanoes 79–80, 79
Armenian Nuclear Power Plant (ANPP)
77–78, 79, 80, 80
likelihood of new volcano formation 85, 85
autocorrelation 61, 66, 69, 130, 132, 134, 134
explosive events 66–69
autocorrelation, partial 61
autocorrelation function (ACF) 61, 134
relationship with frequency spectrum 135, 136
autocovariance 90, 130–137, 176–177
autoregressive (AR) process 61
modelling 132–133, 134
autoregressive–moving average (ARMA) process
ARIMA process 137–138
modelling 133–134
backward shift operator (\(z^{-1}\)) 154
basalt 5–7, 7
Bayes rule/theorem 32, 34
Beta function/distribution 33–34
Big Incomplete Catalogue (BIC) 119, 123, 123
Bingham liquid 262
binomial density function 33
bubble growth 6–7, 6
bulk modulus 208, 222, 225, 262
caldera 40, 79, 81, 116, 262
calibration performance 20–21
Cauchy kernel 81
causal signal
anti-causal 262
non-causal 262
censoring function 53, 53
central limit theorem 119
centralized moving average (CMA) 146
Cerro Negro volcano (Nicaragua)
modelling 1992 eruption 190–191, 190, 191
tephra fallout inversion 235–238, 236, 237, 237
parametric analysis 191–193, 192, 192
Cinder cone 78, 80–81, 80, 98, 190, 235, 262
Clark–Evans test 81, 262
Classical Model for expert judgement 19, 21–23, 22
closure model 201, 262
coda 60, 107, 116, 222–223, 263
coefficient of determination 151–263
colignimbrite plume
(see also co-pyroclastic-flow plume) 263
Colima, Volcán de (Mexico) 58, 58
correlograms of explosive events 67, 68
events per day 64
explosive events data 66
fitted distributions 70
fractal dimension 69
frequency–magnitude plot 68
fumarole pressure 144–145, 144, 145
UC modelling 148–149, 149, 149
magnitude of explosive events 68
repose times distribution 71
study methods 60
characteristics of eruptive events 60
seismic record 60
survival plots 72
correlation coefficient 61, 67, 96, 133, 223, 262
convergent plate margins 1, 2
Confoederatio Helvetica (CH) 60
co-pyroclastic-flow plume 263, 243–245
parametric analysis 191–193, 192, 192, 194
transient modelling of explosive eruptions 202
diffusion of volatiles into bubbles 203
fragmentation criteria 203–204
governing equations 197
mathematical background 185–186, 186
parametric analysis 191–193, 192, 192, 194
transient modelling of explosive eruptions 202
diffusion of volatiles into bubbles 203
fragmentation criteria 203–204
governing equations 202–203
initial and boundary conditions 204
numerical method 204–205
results of numerical simulations 205–206, 205, 205, 206
transient modelling of extrusive eruptions 206
boundary conditions 207–208
governing equations 206–207
numerical method 208
results of numerical simulations 208–210, 209, 210
confidence interval 50–52, 50, 51, 55, 55, 119, 123, 150, 157, 263
correlation coefficient 61, 67, 96, 133, 223
contingency tables 105–107, 112–113
2 × 2 table 106
case study of tornillo at Galeras volcano
(Colombia) 109–112, 110
first mode bands 111
fundamental and higher bands 112
second mode bands 112
coefficients 106
continuous wavelet transform (CWT) 263, 278
corrupt plate margins 1, 2
Cook’s formulation 20
correlation coefficient 61, 67, 96, 133, 223,
238, 263

1750-8207/06/$15.00 © IAVCEI 2006.
Cotopaxi volcano (Ecuador) 116
b value estimates 122
BIC samples 123
confidence intervals for population mean 124
FMR results 119–120, 120
location 116
minimum homogeneity magnitude 121
covariance 134, 135, 146–147, 263
cross covariance 90, 92, 263
cross variogram 90, 91, 94–95, 96–97, 98, 99, 102, 263–264
crust 2, 43, 85, 98, 115, 264
cryptodome 202, 264
crystal growth 6–7, 6, 7
cumulative frequency function 62
cumulative distribution function (CDF) 18, 19, 82, 252, 264, 275
equation 19
data-based mechanistic (DBM) modelling 158–159
data clustering 264
day noise (DN) 162
decision-maker, performance-based (DM) 21–22, 23–24
decompressed jet modelling 186–187, 187
Cerro Negro (Nicaragua) 1992 eruption 190–191, 190
governing equations 197–198
parametric analysis 191–193, 192, 192
deformation of ground, monitoring 10–11
degassing 7–8
monitoring 11
Delphi method 17, 18, 29, 264
delta operator (d) 264
dense rock equivalent (DRE) 180, 249, 253, 264
density function 32–35
destructive plate margins 1, 2
diffusivity 6, 11, 264
Dirichlet distribution 35
discrete Fourier transform (DFT) (see Fourier transform)
discrete wavelet transform (DWT) 164, 166
distribution function (see cumulative distribution function; CDF)
dome (see lava-dome eruption)
downhill simplex method 233–234
dyadic DWT 164, 166, 169, 264, 269, 278
dyke 78, 107, 218, 222–228, 264–265
dynamic autoregression (DAR) model 152–154, 154
dynamic harmonic regression (DHR) model 144, 148, 149, 150–152, 153
dynamic linear regression (DLR) modelling 155–157
dynamics of volcanic explosions 57–58, 69–75
analysis methods
autocorrelation 61
autocorrelation: partial 61
coefficient of variation 60–61
fractal dimension 61–62
time-series analysis 60
data distribution 62
exponential distribution 41–43, 62
gamma distribution 63
log-logistic distribution 43–44, 63
earthquake
log-normal distribution 59, 72, 193, 193, 245
normal distribution 153, 238
Poisson distribution 49, 62
Weibull distribution 62–63
repose interval distribution 69
repose intervals between eruptions 59
study methods 59–60
characteristics of eruptive events 60
study results 63
autocorrelation and fractal dimension analysis 66–69
time-series analysis 63–65
study sites 58
Colima, Volcán de (Mexico) 58
Erebus volcano (Antarctica) 59
Karymsky volcano (Russia) 59
Tungurahua volcano (Ecuador) 59
effusive eruptions 4, 5, 6
basic physical process 3
elastic/elasticity 125, 207–208, 215, 224
epanechnikov kernel 81
epistemic uncertainty 32, 36, 89, 252–255
Erebus, Mount (Antarctica) 58, 59
events per day 65
explosive events data 66
fitted distributions 70
repose times distribution and survival plots 74
study methods 60
characteristics of eruptive events 60
eruption column modelling 185, 194
Cerro Negro (Nicaragua) 1992 eruption 190–191, 190
computational approach 188–189, 188
flowchart 189
governing equations 198–199
mathematical background 186, 186
parametric analysis 191–193, 192, 192, 193
event trees 32, 33
expert judgement 15–30, 265
explosive eruptions 3–5, 4, 5
basic physical process 3
exponential distribution 41–42
explosive events 62
repose intervals between eruptions 41–43
exsolution 3, 5, 73, 181–182, 201, 206–207, 266, 277
extreme value methods 47–48, 48, 54–55
complete series analysis 52–54
return level curves 55
point process model 48–52, 48
likelihood, maximum 51, 53
parameter estimates 50
INDEX

FMR values 119–120, 120
sampling process 119

Hawaiian-style eruption 267, 273, 277
hazard and risk assessment 15
advantages 28
disadvantages 27
Montserrat eruption 24–27
calibration scores for experts 25
probabilistic risk model 27
volcanic event tree 26
opinion elicitation 16
performance-based measures 19–23
quantitative model 31–32, 36
event tree scheme 32–35
long-term assessment example 35–36
posterior density function 32–35
structured expert judgement 17
continuous uncertain quantities 18–19
discrete event probabilities 18
point value estimations 17–18
weighting threshold adjustment 23–24
hazard estimation using geostatistical models 89–90, 101–102
case studies 93
Ostefiel region (Germany) 98–101, 99, 100
Soufrière Hills volcano (Montserrat, West Indies) 96–98, 97, 98, 99
Stromboli volcano (Italy) 93–96, 94, 95
theoretical background
methodology 90–93
origin of geostatistics 90
hazard function
exponential distribution 42
log-logistic distribution 43
hidden Markov models (HMM) 167–168, 167
wavelet-domain 168, 169
hidden Markov tree (HMT) 168
Holocene 2, 39–40, 44–45, 80, 267
hybrid (HY) events 162
ignimbrite 263, 267
informativeness of experts equation 20
input–output modelling 154–155
example—fumarole pressure—temperature
modelling 155–157, 156, 157
integrated random walk (IRW) 146
intensity function 52
integral wavelet transform (IWT) 163–164
isomass map 231, 236–237, 237, 244, 245, 256, 268
isopach map 236, 244, 245, 245, 268
isopleth map 244, 245, 268
term-based weighting 19
Kalman filter (KF) 147–148
Kalman filter–fixed interval smoothing (KF–FIS) algorithms 147–148
Karymsky volcano (Russia) 58, 59
explosive events data 66
fitted distributions 70
study methods 60
characteristics of eruptive events 60

Downloaded from https://pubs.geoscienceworld.org/books/chapter-pdf/4533130/9781862396234_backmatter.pdf on 10 August 2019
INDEX

Monte-Carlo method 100
Monte Carlo method 100
Mooser River (Germany) 99, 101
mother wavelet 163
moving average (MA) 61
moving average 61
see also autoregressive–moving average
(ARMA) process
multiparameter monitoring and modelling of volcanic processes
215–216, 228–229
convolution
seismic experiment on Stromboli (Italy) 216
external triggering 218
data 218–220, 220
model 220–221
gliding spectral lines
seismic observations 221
seismological interpretation 222–223
low-frequency seismic events on Montserrat
(West Indies) 223
magma and conduit parameters 226–227
models 224–226, 225, 226, 227
observational clues 223–224
trigger mechanisms 227–228, 228
removing instrument response 216–218, 217, 219
multiphase (MP) events 162, 267, 271
multiple eruption scenario (MES) 250
multivariate statistics see contingency tables
Nette River (Germany) 99, 101
Nettebach River (Germany) 99, 101
Neuwied (Germany) 99, 101
Newtonian liquid 8, 210, 271
night noise (NN) 162
noise variance ratio (NVR) 147
nonparametric 84, 106, 112, 238, 271
normal distribution 153, 238, 271
Gaussian distribution 81, 82, 118, 129, 177, 232,
266, 271
nucleation 6–7, 6
open-system degassing 7–8
optimization 21–27, 43, 146, 148, 152–153, 271
ordinary differential equations (ODEs), modelling
with 187–188, 189
Osteifel region (Germany) volcanoes 98
hazard estimation
correlation 98–100, 100
data 98, 99
probability maps 101
stochastic forecasting in time 100, 100
stochastic forecasting in time and space, 100–101
parallel computation 270–271
partial autocorrelation function (PACF) 61
partial differential equations (PDEs) 201
Pearson’s contingency coefficient 106
Peclet number (Pe) 202
periodogram 136, 272
permeable foam model 7–8, 8
phreatomagmatic 1, 201, 272

INDEX 283

Pisayambo (Ecuador) seismic zone 116
b value estimates 122
BIC samples 123
confidence intervals for population mean 124
FMR results 120
location 116
minimum homogeneity magnitude 121
Pliocene 80, 272
Plinian eruption column 3
Pleistocene 80, 272
point process modelling 48–52, 48
Poisson distribution 49, 62
explosive events 62
probability density function (pdf) 49
exponential distribution 42
gamma distribution 63
log-logistic distribution 42, 43, 44, 45, 45, 59, 63, 69, 70, 72–74
log-normal distribution 59, 72, 193, 193, 245
normal distribution 153, 238
Poisson distribution 49, 62
Weibull distribution 62–63
probability of discrete events 18
probability of explosive eruptions 39, 44–45
extinction versus dormancy 39–40
record of volcanic activity 40–41
repose intervals 41
exponential distribution 41–43
log-logistic distribution 43–44
proper scoring rule 22, 262, 272
pumice 182, 272, 276
pyroclast 261, 262, 268, 272
pyroclastic flow 3, 5
quality factor 107, 272
Quaternary 39, 79–84, 79, 89, 93, 190, 273
Raleigh–Lamb equation 203
random walk (RW) 146
rarefaction wave 202, 205, 273
recursive fixed interval smoothing 273
recursive parameter estimation 273
recursive state estimation 273
refined instrumental variable (RIV) algorithm 155
repose interval 40–45, 41, 42, 43, 44, 44, 57, 59, 62–63, 64, 65, 70–71, 74, 175, 180, 182, 273
reticulate 273
Reventador volcano (Ecuador) 116
b value estimates 122
BIC samples 123
confidence intervals for population mean 124
FMR results 119–120, 120
location 116
minimum homogeneity magnitude 121
Reynolds number 188, 196, 233, 241, 245–246, 273
Rhein River (Germany) 99, 101
rheology of magma 8–9, 9
rhyolite 3, 6, 181, 273, 274
Rieden (Germany) 99, 101
‘Ring of Fire’ 2
root-mean-square error (rms) 273
Runge–Kutta 187, 189, 274
saturation 6, 171
scoria 262, 274, 276
scoring rule 22, 272, 274
seismological data, monitoring 10
Sevan, Lake (Armenia) 79
Shamiram Plateau (Armenia) 79
short-time Fourier transform (STFT)
seismic signals of volcanic origin (SSVO) classification 161, 172
dataset from Mount Merapi (Indonesia) 161–163, 162
Fourier transform and fast wavelet transform 163–164, 165, 166, 167
Markov models 164–168
wavelet-domain 168, 169
pattern recognition system 168
classification scheme evaluation 171–172
classifier selection 169–171, 170
confusion matrices 171
feature extraction 168–169
feature selection 169
theoretical background 163
silicic 57, 79, 202, 268–269, 274
sill 274
Small Incomplete Catalogue (SIC) 119, 123
Soufrière Hills volcano (Montserrat, West Indies) 24–27, 96, 175, 176, 182
calibration scores for experts 25
explosion sequence 176
hazard estimation 96–97, 98
data 96, 97
identification of precursors 97–98, 99
model development 179–182
schematic representation 180
multiparameter monitoring and modelling 223
low-frequency event models 224–226, 225, 226, 227
magma and conduit parameters 226–227
observational clues 223–224
trigger mechanisms 227–228, 228
probabilistic risk model 27
stochastic analysis methods 175–177
results 177–179
variograms 178, 179
volcanic event tree 26
spatial distribution of volcanoes 77–78, 84–85
application of Gaussian and adaptive Gaussian kernels 80
contour plots of volcano density 83
cumulative density distribution 82
defining events 80–81
modelling spatial intensity 81–84
Armenia 79–80, 79
nature of distribution in space and time 78–79
rate of volcanic activity 84
probability estimates 84
spectral lines, gliding 221
speedup 233
standard error 42
state-dependent parameter (SDP) estimation, 157–158, 158
...
strain 8–9, 9, 227–228, 275
stress 8–10, 78, 125, 140, 186, 201, 216–220
Stromboli-style eruption 59–60, 191, 217, 262, 275
Strombolian-style eruption 59–60, 191, 217, 262, 275
Stromboli volcano (Italy) 93
hazard estimation correlation 94–95
data 93–94, 94
identification of precursors 95–96, 95
multiparameter monitoring 216
supersaturation 3, 6, 262, 275
survivor function exponential distribution 42, 42, 43
log-logistic distribution 43
tephra dispersion, probabilistic modelling 243, 256–257
challenges and hazards 243–244
considerations
choice of activity scenarios 255–256
computing implications 256
importance of field data 255
importance of probabilistic analysis 254–255
uncertainty analysis 255
field investigations 244–245
contour maps 244
numerical investigations 245–246
rectangular coordinate system 246
probabilistic analysis 246–247
accuracy of Monte-Carlo technique 251
hazard curves 252–253, 252, 253
input parameters 254
maps based on hazardous thresholds 247–252, 248, 249
multiple eruption scenario (MES) 250
probability maps and hazard curves based on return periods 253–254, 253
assessing inversion performance 238–240, 239
Cerro Negro (Nicaragua) 1992 eruption 235–238, 236, 237, 237
downhill simplex method 233–234
numerical simulation 232–233, 233
tidal potential 218, 275
tilt cycles, correlation with seismicity 223–224
time-series 129, 141–142, 159
data-based mechanistic (DBM) modelling 158–159
input–output modelling 154–155
example–fumarole pressure–temperature modelling 155–157, 156, 157
model-based systems for analysis and forecasting 143–144
new approaches to volcanic series 143
non-stationary 137
example–medical time series 140–141, 140, 141
example–seismic time series 139–140, 139
global non-stationarity 137–138
local non-stationarity 138
testing for non-stationarity 137
other aspects 141
discrete-valued time series 141
forecasting 141
multivariate systems 141
umbrella cloud 3
uncertain quantities 18–19
unobserved component (UC) modelling 144–145
examples 148
Colima volcano pressure 148–149, 149, 149
seismic time series 149–154, 150, 151
fumarole pressure 144, 145
state estimation 147–148
theoretical and algorithmic background 145–146
variation, coefficient of 60–61
variograms 90–91
very-long-period earthquake (VLP)  (see earthquake)
vesicle 197, 276
volcanic explosivity index (VEI) 32, 34–35, 40
hazard rates
log-logistic distribution 45
likelihood, maximum exponential distribution 42
log-logistic distribution 44
record over time 40–41, 41, 41
survivor function
log-logistic distribution 44
volcanic processes 1, 12, 15–16
data from monitoring 10
gas monitoring 11
ground deformation 10–11
seismic monitoring 10
explosive versus effusive activity 2–3
effusive eruptions 3, 4, 5
effusive eruptions 3–5, 3, 4
lava-dome eruptions 3, 5
modelling approaches 11–12
pre-eruptive physical processes 5–6
  fragmentation and gas–particle flow regime 9–10, 9
magma rheology 8–9, 9
nucleation and growth of bubbles and crystals 6–7, 6, 7
open-system degassing 7–8
risk from 1, 15
spatial and temporal distributions 1–2
  distribution relative to tectonic plates 2
historical record 2
volcanic tremor 10, 93, 95, 107, 143, 216, 221, 227
volcano-tectonic events, type B (VTB) 162

INDEX

Vulcanian-style eruption 59, 60, 63, 71, 73, 90, 96, 175, 176, 179, 182, 244, 249, 253, 277
vulnerability 32, 89, 277
wavelet
  wavelet transform 163–164, 168, 168, 170, 172, 277–278
Wehr (Germany) 99, 101
Weibull distribution for explosive events 62–63
weighting of individual experts equation 21
Yerevan (Armenia), 79
Z-transform 278