BIOSENOSRS

BIO 580

Error analysis in biosensing

WEEK-14

Fall Semester

Faculty: Dr. Javed H. Niazi KM
Faculty of Engineering & Natural Sciences
Sabanci University

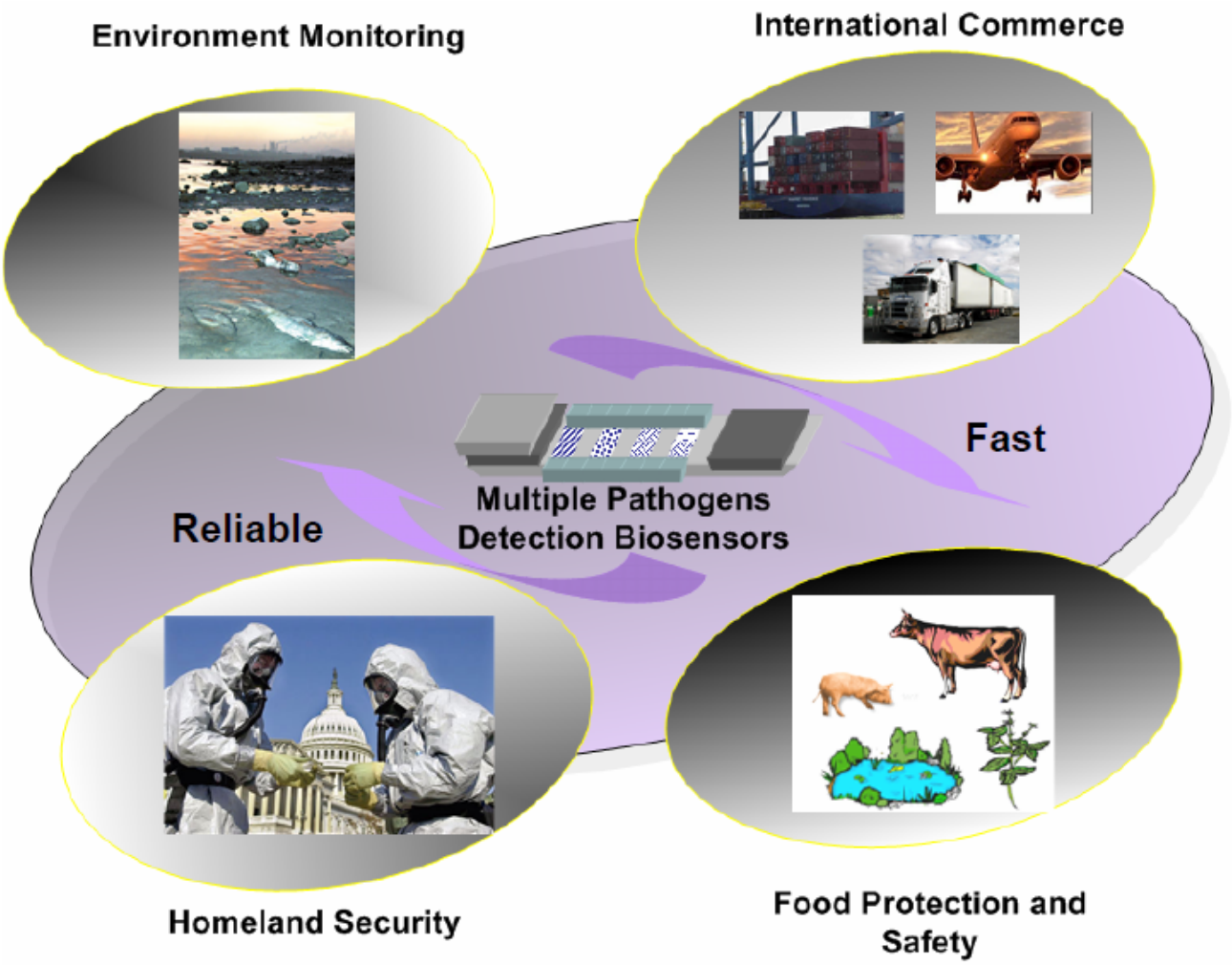
Topics that will be covered in the course

- ❑ History of biosensor development, applications and requirements of biosensors and classification
- ❑ Principles of molecular recognition and transduction signal acquisition
  - ✓Sources of Biological Recognition elements – enzymes/proteins, ssDNAs, antibody and Others
  - ✓Design considerations for use of recognition elements in biosensors
  - ✓Modeling of reactions for various biosensor applications- electrochemical, optical, piezoelectric, colorimetric, fluorometric and others.
- ❑ Modification of sensor surfaces and immobilization techniques
  - ✓Covalent modification of surfaces using surface chemistry
  - ✓Self Assembled Monolayers (SAM) and adsorptions
  - ✓Other ways to immobilize biological macromolecules on various solid surfaces
- ❑ Detection methods and Physical Sensors
  - ✓Electrodes/transducers – electrochemical (amperometric, potentiometric, and conductimetric transductions)
  - ✓Other sensors - for e.g., optical sensors (colorimetric/fluorimetric/luminometric sensors), Surface Plasmon Resonance (SPR) sensors, and piezoelectric resonators.
- ❑ Fabrication of biosensors
  - ✓Miniaturization-application of nano-materials, nanoparticles, carbon nanotubes (CNTs) and others
  - ✓Biocompatibility – stability, reproducibility and repeatability of biomolecules on transducer surfaces
- ❑ Data acquisition, statistical and error analysis
  - ✓Inter and Intra-assays and Coefficient of variation (CV)
  - ✓Signal to noise ratio
  - ✓Normalization/optimization and signal retrieval
- ❑ Examples of commercial biosensors

Design Parameters for Selecting a Transducer for a Biosensor

Parameter	Definition
Sensitivity	Ratio of change of sensor output, Δy, to change in the input, Δx: $S = \frac{\Delta x}{\Delta y}$. Maximum sensitivity is normally a design goal as long as other key parameters, such as linearity and accuracy, are not sacrificed.
Linearity	The constancy of the ratio of output to input. If relationship is linear then $y = Kx$ where K (also called gain) is a constant. If the sensor is nonlinear then the input–output relationship is given by an equation: $y = f(x)$.
Working range	Difference between the maximum and minimum values that can be measured by the biosensor or sensor element. A large working range is often preferred.
Accuracy	Difference between the measured and the actual values. It can either be defined as an absolute value which is the maximum error within the working range, or a relative value which is the ratio of the maximum error to the range of measurement.
Repeatability	Difference in value between two successive measurements under the same operating environment.
Resolution	Minimum change of input that can be detected at the output of the biosensor. In some circumstances it is preferred to define resolution as the number of incremental measurements within the range from minimum to maximum value.
Output	Voltage signal is preferred because microcontrollers and computers are being used to automatically gather the data. If the proposed biosensor output signal is not a voltage then it must be converted to an electrical signal with additional hardware, increasing the source of errors.
Response time	Time required for a change in the input to be observed as a stable output change (also known as *settling time*). In most cases, the response time is measured from the start of an input change to the time when the output has settled to the specified range. The response time of the sensor system is determined by its time constant.
Bandwidth	Under ideal conditions, a transducer would produce identical amplitude responses for input signals with the same amplitude, independent of their constituent frequencies. However, the output of a physical transducer is dependent on the amplitude and frequency of an input signal because of the transducer's inherent bandwidth limitations. This characteristic is often described as the transfer function in the frequency domain.

Motivation

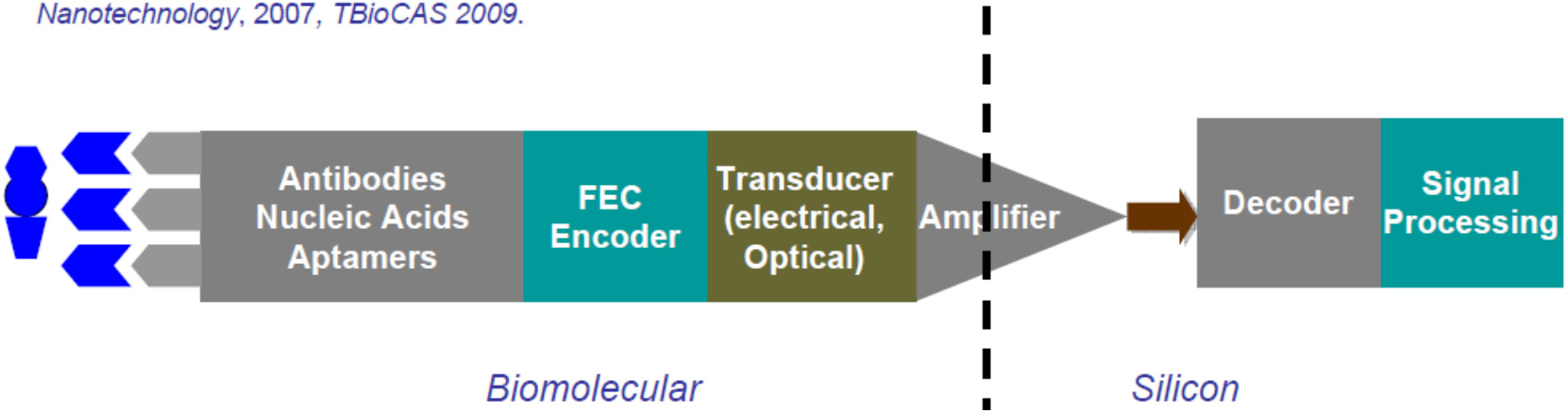


Environmental variability and non-deterministic response of bio-molecules affects reliability of biosensors.

Can forward error correction (FEC) principles be used to improve reliability of biosensors ?

Forward error-correction biochip

Nanotechnology, 2007, *TBioCAS* 2009.

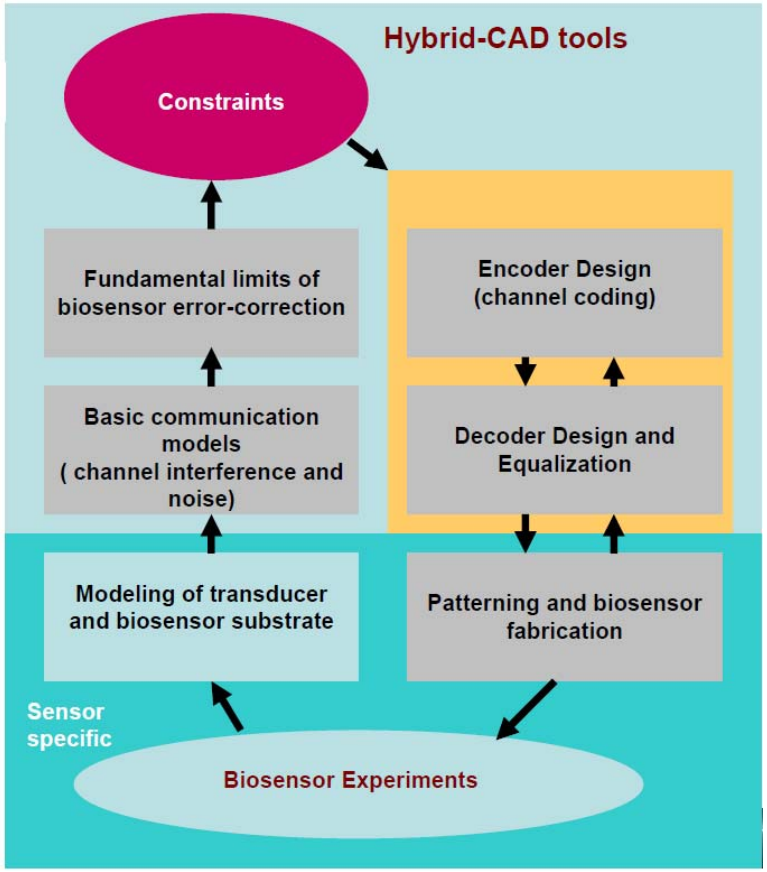


The "Big" Picture

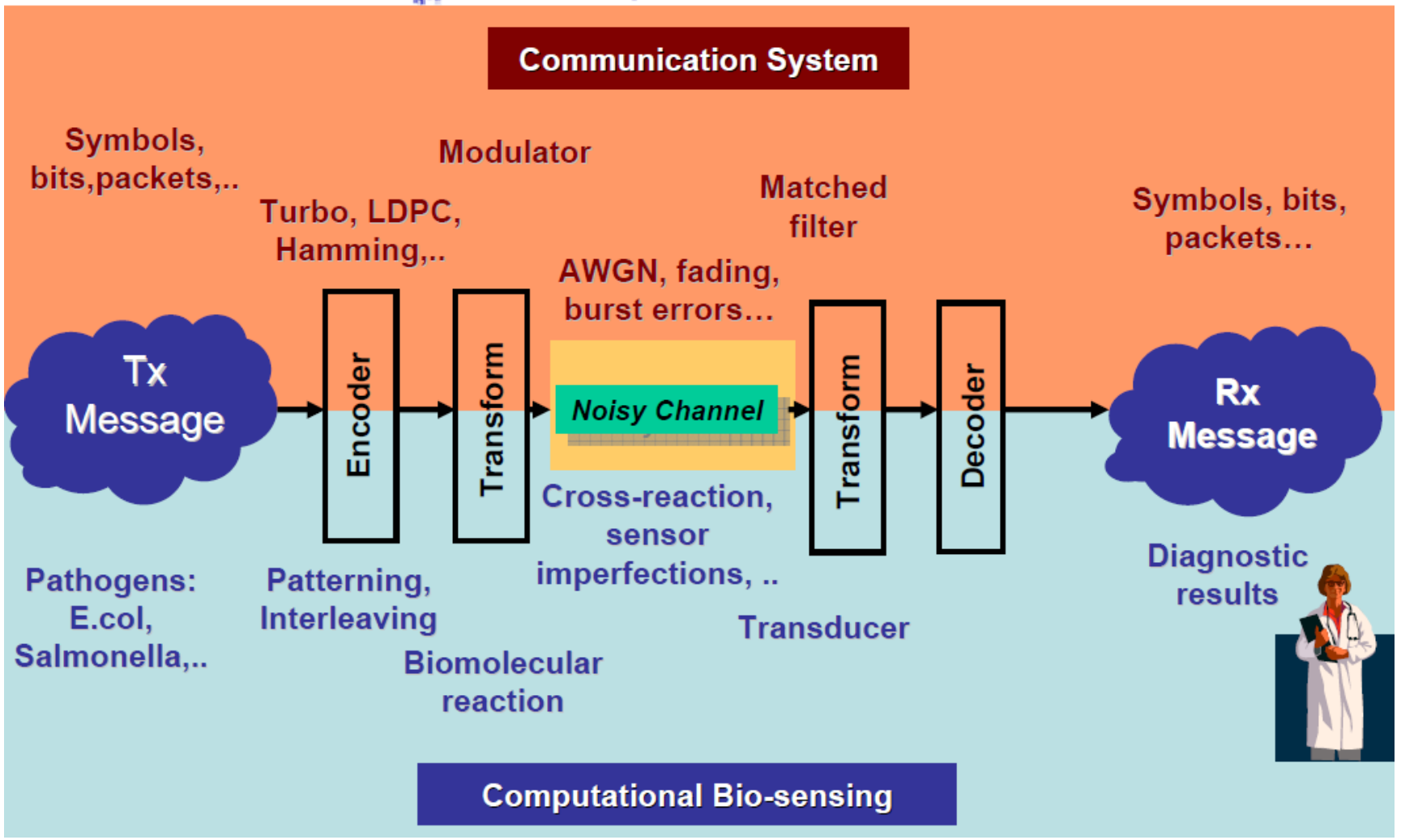
Virtual biosensor: Alleviate cumbersome experimental cycles and protocols.

For a given substrate of fixed size how many pathogens can be detected reliably at a specified error rate ?

Can new encoding-decoding principles be discovered by using CAD models ?

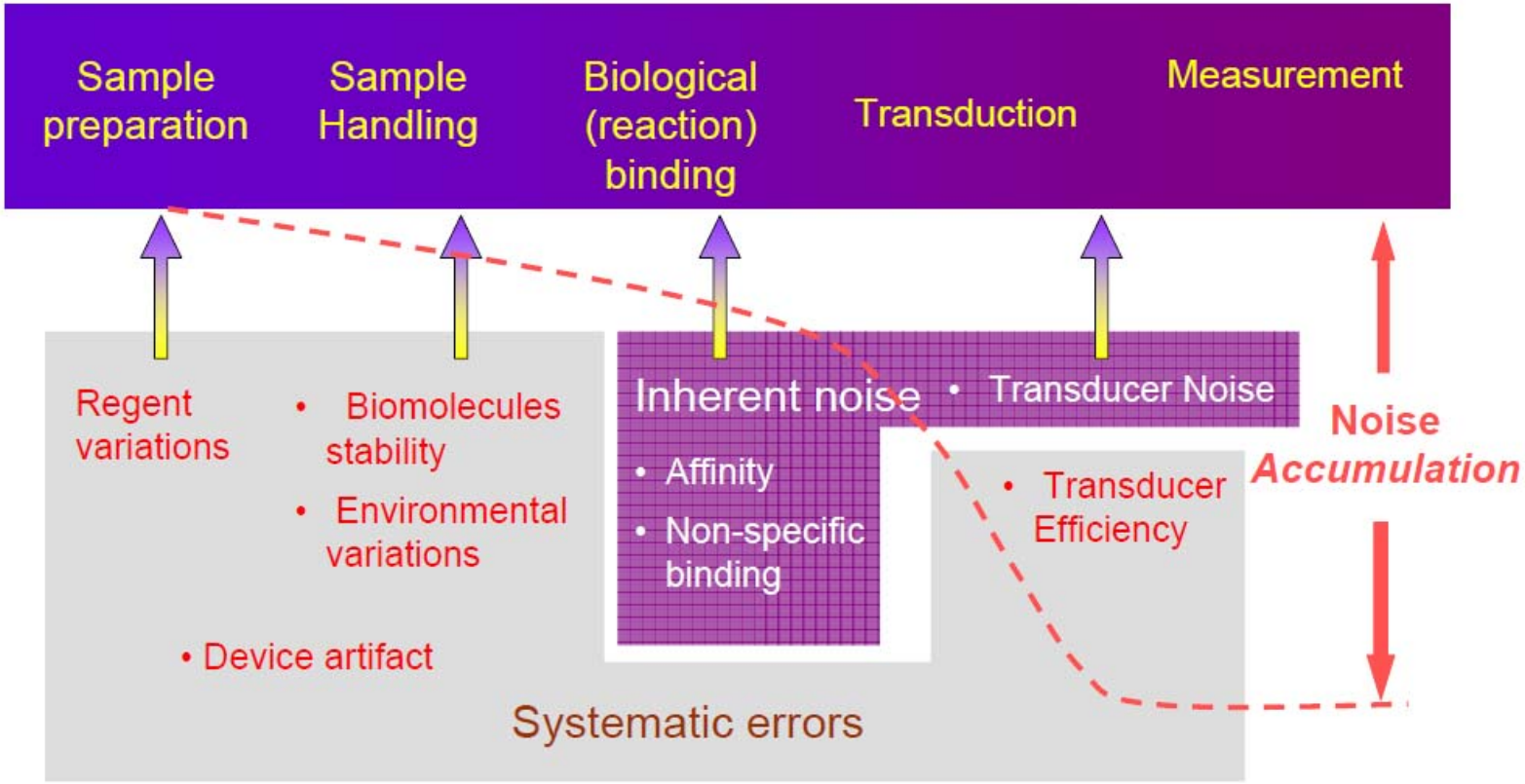


Communication model for reliable bio-sensing



Lessons Learned from the Past

- Signal Amplification at the bio-molecular level significantly improves the sensitivity of the biosensor.



Key References

Y. Liu, S. Chakrabartty*, and E. C.Alocilja, "Fundamental Building Blocks for Molecular Bio-wire based Forward-error Correcting Biosensors", Nanotechnology, 18, (2007), 4240172.

Y. Liu, A. Gore, S. Chakrabartty*, and E. C.Alocilja, "Characterization of Sub-systems of a Molecular Bio-wire based Biosensor Device," Microchimica Acta , 2008, DOI: 10.1007/s00604-008-0950-0.

Y.Liu, S.Chakrabartty*, Factor Graph based Biomolecular Circuit Analysis for Designing Forward Error Correcting Biosensors, IEEE Transactions of Biomedical Circuits and Systems, vol. 3, no. 3, pp.150-159, June 2009

Google book for mathematical modeling and error analysis

Relative standard deviation

In probability theory and statistics, the **relative standard deviation (RSD or %RSD)** is the absolute value of the coefficient of variation. It is often expressed as a percentage. A similar term that is sometimes used is the **relative variance** which is the square of the coefficient of variation. Also, the **relative standard error** is a measure of a statistical estimate's reliability obtained by dividing the standard error by the estimate; then multiplied by 100 to be expressed as a percent
The relative standard deviation is widely used in analytical chemistry to express the precision and repeatability of an assay.

(standard deviation of array X) × 100 / (average of array X) = relative standard deviation

Coefficient of variation (CV)

In probability theory and statistics, the **coefficient of variation (CV)** is a normalized measure of dispersion of a probability distribution. It is defined as the ratio of the standard deviation σ to the mean μ :

$$c_v = \frac{\sigma}{\mu}$$

This is only defined for *non-zero* mean, and is most useful for variables that are always *positive*. It is also known as **unitized risk.**

The coefficient of variation should only be computed for data measured on a ratio scale. As an example, if a group of temperatures are analyzed, the standard deviation does not depend on whether the Kelvin or Celsius scale is used. However the mean temperature of the data set would differ in each scale and thus the coefficient of variation would differ. So the coefficient of variation does not have any meaning for data on an interval scale.
Standardized moments are similar ratios, μ^k/σ^k, which are also dimensionless and scale invariant. The variance-to-mean ratio, σ^2/μ, is another similar ratio, but is not dimensionless, and hence not scale invariant.

Normalization (statistics) for further ratios.
In signal processing, particularly image processing, the reciprocal ratio μ/σ is referred to as the signal to noise ratio.

Advantages

The coefficient of variation is useful because the standard deviation of data must always be understood in the context of the mean of the data. The coefficient of variation is a dimensionless number. So when comparing between data sets with different units or widely different means, one should use the coefficient of variation for comparison instead of the standard deviation.

Response time

This response is characteristic of well-damped, inherent second-order systems. A single parameter from the response curve is usually reported, such as T95, the 95% response time, defined as the time required for the amplitude of the signal to change from its initial steady state value to 95% of the final steady state value.

Dynamic Delay and Dynamic Error

The dynamic delay has been used historically to describe the delay, lag, or temporal displacement between physical transients and the sensor responses to those transients.

The dynamic delay is specified solely by properties of the biosensor and external mass transfer.

The dynamic error, a related term, is the difference between the actual value of the variable at a given moment and the value simultaneously reported by the sensor.

The dynamic error is the product of the dynamic delay and the instantaneous rate of analyte change.

The dynamic delay and dynamic error have been adapted for in vitro characterization of biosensors exposed to concentration challenges in the form of linear ramps

Signal-to-noise ratio (often abbreviated **SNR** or **S/N**) is an electrical engineering measurement, also used in other fields (such as scientific measurement or biological cell signaling), defined as the ratio of a signal power to the noise power corrupting the signal. A ratio higher than 1:1 indicates more signal than noise.

In less technical terms, signal-to-noise ratio compares the level of a desired signal (such as music) to the level of background noise. The higher the ratio, the less obtrusive the background noise is.

Technical sense [edit]

In engineering, signal-to-noise ratio is a term for the power ratio between a signal (meaningful information) and the background noise:

$$\mathrm{SNR} = \frac{P_{\mathrm{signal}}}{P_{\mathrm{noise}}},$$

where P is average power. Both signal and noise power must be measured at the same or equivalent points in a system, and within the same system bandwidth. If the signal and the noise are measured across the same impedance, then the SNR can be obtained by calculating the square of the amplitude ratio:

$$\mathrm{SNR} = \frac{P_{\mathrm{signal}}}{P_{\mathrm{noise}}} = \left(\frac{A_{\mathrm{signal}}}{A_{\mathrm{noise}}}\right)^2,$$

where A is root mean square (RMS) amplitude (for example, typically, RMS voltage). Because many signals have a very wide dynamic range, SNRs are usually expressed in terms of the logarithmic decibel scale. In decibels, the SNR is, by definition, 10 times the logarithm of the power ratio:

$$\mathrm{SNR(dB)} = 10\log_{10}\left(\frac{P_{\mathrm{signal}}}{P_{\mathrm{noise}}}\right) = 20\log_{10}\left(\frac{A_{\mathrm{signal}}}{A_{\mathrm{noise}}}\right) = P_{\mathrm{signal,dB}} - P_{\mathrm{noise,dB}}.$$

Alternate forms [edit]

A common alternative definition of SNR is the ratio of mean to standard deviation of a signal or measurement:[1][2]

$$\mathrm{SNR} = \mu / \sigma$$

where μ is the signal, or the mean or expected value of the signal, or some measure of signal strength, and σ is the standard deviation of the noise, or an estimate thereof. The exact methods may vary between fields. For example, if the signal data are known to be constant, then σ can be calculated using the standard deviation of the signal. If the signal data are not constant, then σ can be calculated from data where the signal is zero or relatively constant.

v · d · e	Statistics		[hide]
Descriptive statistics	**Continuous data**	**Location**	Mean (Arithmetic, Geometric, Harmonic) · Median · Mode
		Dispersion	Range · Standard deviation · **Coefficient of variation** · Percentile
		Moments	Variance · Semivariance · Skewness · Kurtosis
	Categorical data		Frequency · Contingency table
Inferential statistics and hypothesis testing	**Inference**		Confidence interval (Frequentist inference) · Credible interval (Bayesian inference) · Significance · Meta-analysis
	Design of experiments		Population · Sampling · Stratified sampling · Replication · Blocking · Sensitivity and specificity · Optimal design
	Sample size estimation		Statistical power · Effect size · Standard error
	General estimation		Bayesian estimator · Maximum likelihood · Method of moments · Minimum distance · Maximum spacing
	Specific tests		Z-test (normal) · Student's t-test · F-test · Chi-square test · Pearson's chi-square test · Wald test · Mann–Whitney U · Wilcoxon signed-rank test
	Survival analysis		Survival function · Kaplan–Meier · Logrank test · Failure rate · Proportional hazards models
Correlation and regression analysis	**Correlation**		Pearson product-moment correlation · Rank correlation (Spearman's rho, Kendall's tau) · Confounding variable
	Linear regression		Simple linear regression · Ordinary least squares · General linear model · Analysis of variance · Analysis of covariance
	Non-standard		Nonlinear regression · Nonparametric · Semiparametric · Robust
	Non-normal errors		Generalized linear model · Binomial · Poisson · Logistic
Statistical graphics	Bar chart · Biplot · Box plot · Control chart · Correlogram · Forest plot · Histogram · Q-Q plot · Run chart · Scatter plot · Stemplot		
Category · Portal · Outline · Index			

Assay Quality Control

- Results produced by a given laboratory for each assay are consistent over time and that results produced by different laboratories from the same samples are comparable.

- One published study has investigated within- and between-laboratory variation in Elisa results in the USA (Kreider 1991b). Results varied significantly and substantially among different laboratories (the greater source of variation) and among different days in the same laboratory.

- This suggests that single determinations on individual serum samples are not likely to give a reliable estimate of antibody titre. The large variability within laboratories further indicates the need for standard reference pools of positive serum to be included in assays in order to substantiate assay results.

- Murray et al. (1993) have discussed in some detail the sources of variability of assay results. While these authors concentrated on biochemical assays the same basic principles apply equally well to serology. Each assay must be validated to identify and quantify sources of variation in results. We must also keep in mind that there are non-assay sources of variation. These can be grouped into those factors which precede the assay (how and when a sample is taken, how it is manipulated, stored, transported, and identified) and those which come after (for example transcription errors in report generation).

In validating an assay the following areas need to be addressed

1. **Specificity** - A highly specific assay will have a low tendency to show "false positive" reactions on sensor surface exposed to a closely related target molecule. This can be tested by obtaining mono-specific sera raised against a range of other targets and including them in the test.

2. **Sensitivity** - This is a measure of the ability to detect clinically important but very low levels of antibody/antigen. There will sometimes be a trade-off in that the higher the sensitivity of an assay the lower its specificity. The sensitivity of a test can be evaluated by diluting known positive samples sequentially and determining the dilution at which the reaction is lost.

3. **Accuracy** - This is a measure of the ability of the test to measure purified amounts of the substance sought when it is added in measured amounts to a typical test sample. Rarely will we have purified antibody available for this type of study but neither is this required in that we will not be reporting results in "milligrams of antibody". If known positive field sera are available then they can be used as a pool in repeated assays. This will be most valuable if this pool is also submitted to a reference laboratory for testing using a recognized and already validated procedure. The alternative is to take purified mono-specific antiserum and use this to "spike" sero-negative field serum at different concentrations then use the spiked samples to establish a measure of the variability of the assay results within a given sample.

4. **Precision** - This is the ability of the assay to consistently reproduce a result when sub-samples are taken from the same specimen. Within-assay and inter-assay precision are two distinct measures of this can be made as part of the validation procedure. The formulae used for the calculation of CV% are slightly different from the conventional formula (Standard Deviation divided by the mean and multiplied by 100).

Within-assay precision - Assay 10 duplicated samples on the same plate (a total of 20 assays) and calculate an intra-assay coefficient of variation as follows:

$$\frac{\text{Mean of the Standard Deviations of the Duplicates}}{\text{Grand mean of the duplicates}} \times 100$$

A figure of 10% or less is considered satisfactory (Murray et al 1993)
Inter-assay Precision

In this case the 10 runs on duplicate samples are run on different days. For each run the mean, the deviation, and the % C.E. are calculated. The interassay coefficient of variation is calculated from the formula:

$$\frac{\text{Standard Deviation of the means of the duplicates}}{\text{Grand Mean of the Duplicates}} \times 100$$

This system should include a definition of criteria for assay acceptability, along with a means for identifying sources of variation and implementing corrective procedures. One component of such a system might be to repeat an Inter-assay Precision test. Alternatively, simply calculate the %CV for all assays of each standard control serum each month and plot the results over time. A reasonable target for %CV in routine testing is 10-15%.