Computer Aided Process Engineering CL303

Module-4

Probability distribution functions in engineering application and its statistics

Arnab Karmakar

BIT Mesra, Ranchi

Probability distributions of continuous variables.
Probability density function

$$Pr[a < x < b] = \int_a^b f(x)\, dx.$$

$f(x)$ = Probability density function

$$A = Pr[a<x<b] = \int_a^b f(x)\, dx.$$

$$Pr[x < x_1] = \int_{-\infty}^{x_1} f(x)\, dx$$

$$\sum p(x_i) \sim \int f(x)\, dx.$$

Discrete   continuous.

$$\therefore \int_{-\infty}^{+\infty} f(x)\, dx = 1$$

normal distribution

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Binomial distribution

For only two possible outcomes : head or tail, success or failure, defective item or good item.

Let the probability that an item is defective be p. so the probability that an item is good be q. So $p+q=1$.

Let the fixed no. of trial be n.

Then the general expression for the probability of exactly r defective items in any order in n trial

$$Pr[R=r] = {}^nC_r \, p^r q^{n-r}.$$

$$= \frac{n!}{r!(n-r)!} p^r \cdot q^{n-r}$$

Problem A company is considering drilling. four oil ~~oil~~ wells. Probability of success for each well is 0.4, indedendent of the results for any other well. The cost of each well is $200,000. Each well that is successful will be worth $600,000.

a) Determine probability distribution.
b) What is expected number of success.?
c) What is expected gain.?
d). What is gain if only one well is successful?
e). What is probability of a loss.?
f). What is standard deviation?

Varience

$$\sum_{i=1}^{N} (x_i - \mu)^2 / N$$

$$\sigma_x^2 = E(x-\mu_x)^2 = \text{mean } (x-\mu_x)^2$$

$$= \sum (x-\mu_x)^2 Pr(x_i).$$

$$\sigma_x^2 = \Sigma[x^2] - \mu_x^2$$

$$= E[x^2 - 2\mu_x X + \mu_x^2]$$

$$= E[x^2] - 2\mu_x E[X] + \mu_x^2$$

$$= E[x^2] - 2\mu_x^2 + \mu_x^2$$

$$= E[x^2] - \mu_x^2$$

$$E[x^2] = \sum x_i^2 Pr(x_i).$$

standard deviation $\sqrt{\sigma_x^2} = \sigma_x.$

$$E(R^2] = 0^2 * \frac{1}{32} + (1)^2 \frac{5}{32} + 2^2 \frac{10}{32} + 3^2 \frac{10}{32} + 4^2 \frac{5}{32} + 5^2 \frac{1}{32}$$

$$= 7.5.$$

$$\therefore \sigma_x^2 = 7.5 - (\mu_R)^2 = 7.5 - (2.5)^2 = 1.25$$

Std. $\sigma_x = \sqrt{1.25} = 1.118$

Poisson Distribution (S. D Poisson, French Mathemati

↳ counts from Geiger counter.
collisions of cars at specific intersection under specific conditions,
flaws in casting

Probability of r occurrences in a fixed interval of time or space under particular condition. is given by.

$$Pr[R=r] = \frac{(\lambda t)^r e^{-\lambda t}}{r!}$$

t = interval of time or space in which events occur.
λ = mean rate of occurrence per unit time or space.
λt = dimensionless

$$Pr[R=r+1] = \left(\frac{\lambda t}{r+1}\right) Pr[R=r]$$

Prob
The number of meteors found by a radar system in any 30-sec interval under specified conditions averages 1.81. Assume it appear randomly and independently.

a) What is the probability that no meteors are found in a one-minute interval?

b) Find. $Pr[8 \geq r \geq 5]$.

a) $\lambda = 1.81/0.5 \text{ min}^{-1} = 3.62 \text{ min}^{-1}$.

$\therefore t = 1; \mu = \lambda t = 3.62 \times 1 = 3.62$.

$Pr[r=0] = e^{-3.62} = 0.0268$

b) $\mu = \lambda t = 3.62 \times 2 = 7.24$ occurance in 2 minute.

$$Pr[R=5] = \frac{(7.24)^5 e^{-7.24}}{5!} = 0.1189$$

$$Pr[R=r+1] = \frac{\lambda t}{r+1} Pr[R=r]$$

$$Pr[R=6] = \frac{7.24}{6} 0.1189 = 0.1435$$

$$Pr[R=7] = \frac{7.24}{7} \times 0.1435 = 0.1484$$

$$Pr[R=8] = \frac{7.24}{8} \times 0.1484 = 0.1343$$

$$Pr[5 \leq r \leq 8] = 0.1189 + 0.1435 + 0.1484 + 0.1343$$

$$= \underline{\underline{0.545}}$$

Probability → It is a measure of the likelihood that a particular event occur.

Probability distribution of discrete variable

Probability functions

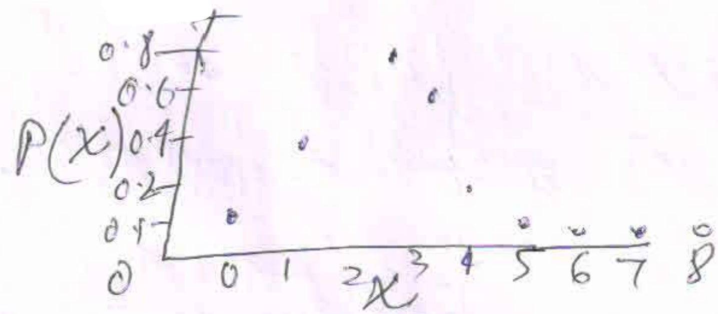
discrete random variable X are

$x_0 \; x_1 \; x_2 \ldots \ldots x_K$.

Probabilities $P(x_0) \; P(x_1) \; P(x_2) \ldots P(x_K)$.

$$P(x_i) \geq 0 \quad \text{and} \quad \sum_{i=0}^{K} P(x_i) = 1.$$

$P(x_i)$ = probability function



cumulative distribution function

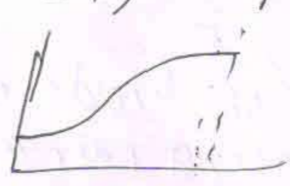
$$Pr[X \leq x] = \sum p(x_i).$$

$$Pr[X \leq 3] = p(x_0) + p(x_1) + p(x_2) + p(x_3)$$
$$= p(0) + p(1) + p(2) + p(3).$$

$$Pr[X \leq 2] = p(0) + p(1) + p(2).$$

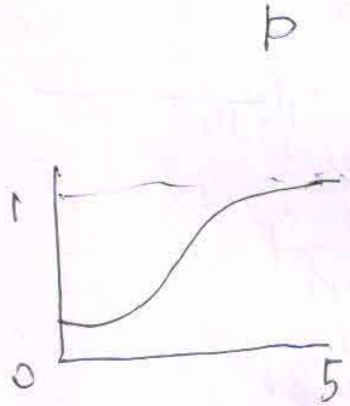
$$p(3) = Pr[X \leq 3] - Pr[X \leq 2].$$

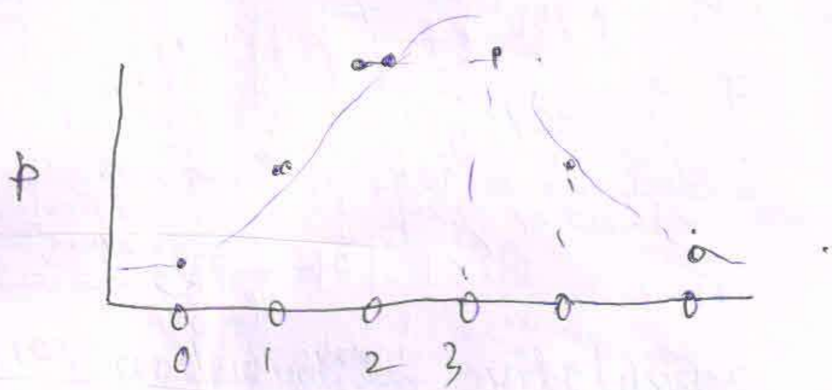
$$p(x_i) = Pr[X < x_i] - Pr[X \leq x_{i-1}]$$



Five fair coin

no. of heads (r)	p(r)	
0	$\frac{1}{32}$	${}^5C_0 (0.5)^0 . (0.5)^5$
1	$5/32$	$\frac{5!}{5!\,0!} (0.5)^5$
2	$\frac{10}{32}$	
3	$\frac{10}{32}$	
4	$5/32$	
5	$1/32$	





$$p(3) = Pr[R \le 3] - Pr[R \le 2]$$

$$= \frac{26}{32} - \frac{16}{32} = 0.3125$$

$$P_r[x_i] = \frac{f(x_i)}{\Sigma f(x_i)}$$

Mean

$$E(x) = \mu_x = \Sigma x_i P_r[x_i]$$

$$E(R) = \mu_R = 0 \times \frac{1}{32} + 1 \times \frac{5}{32} + 2 \times \frac{10}{32} + 3 \times \frac{10}{32} + 4 \times \frac{5}{32} + 5 \times \frac{1}{32}$$

$= 2.5$ no. of heads obtained on tossing fair coin.

f) $\sigma_x^2 = E(x^2) - \mu_x^2$

$E(X)^2 = 3.5200 \cdot \quad \mu_x = 1.6 .$

$\therefore \sigma_x^2 = 3.5 - (1.6)^2 = 0.96$

$\sigma_x = \sqrt{0.96} = \underline{\underline{0.98}}$

Reference : W. J. DeCoursey, Statistics and Probability for Engineering Applications.

Stochastic Process

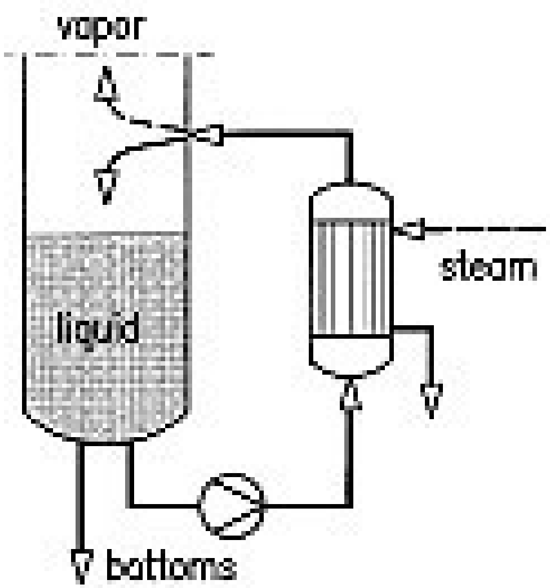
Probability distribution functions in engineering applications and their statistics

Probability density functions in Engineering Application

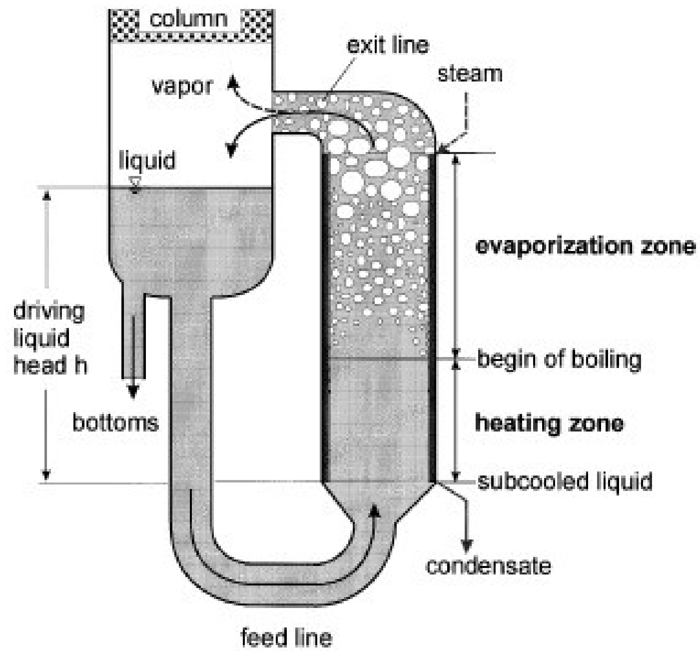
- Flow regime and pattern identifications in boiling channel to predict heat transfer and ensure stable, safe operation of evaporator, drum type boiler, reboiler, nuclear reactor cooling application, exothermic reactor cooling, etc.
- Heat transfer coefficient in the boiling channel is the function of properties of the fluid, flowrate, temperature difference, quality/void fraction, and flow regime of boiling flow.
- Accurate estimation of heat transfer coefficient and flowrate requires the knowledge of flow regimes and void fraction.

NATURAL & FORCED CIRCULATION BOILING- APPLICATIONS

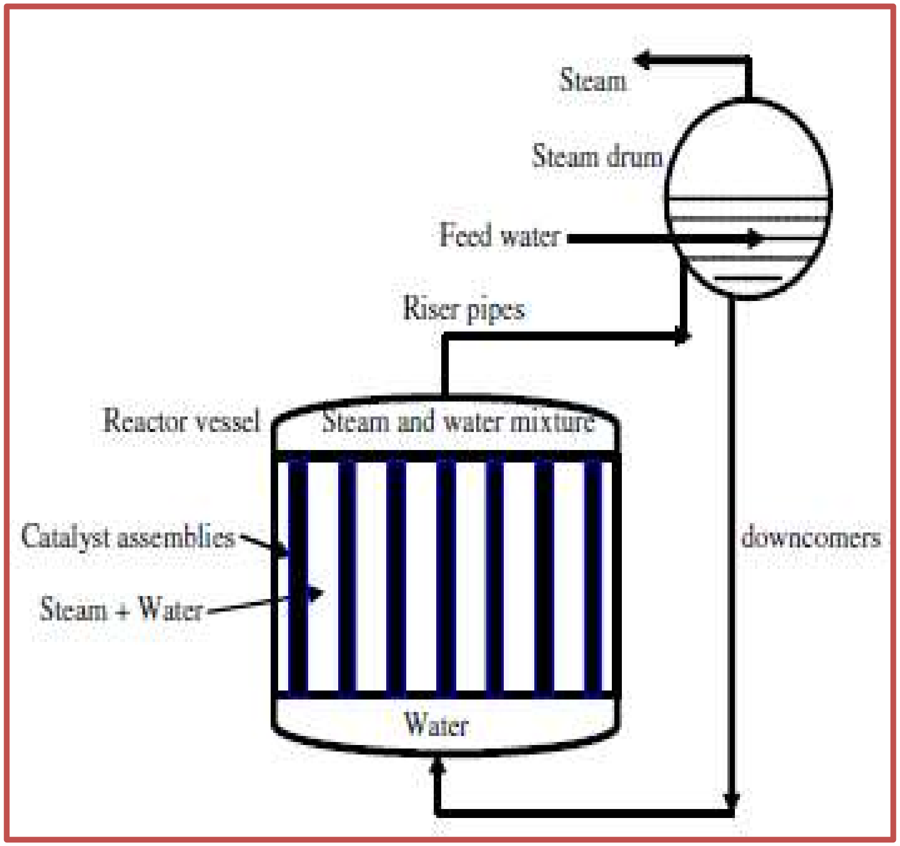
Forced circulation boiling channel is an integral part of forced circulation boiling loop employed in evaporator, drum type boiler, reboiler, nuclear reactor cooling application, exothermic reactor cooling etc .

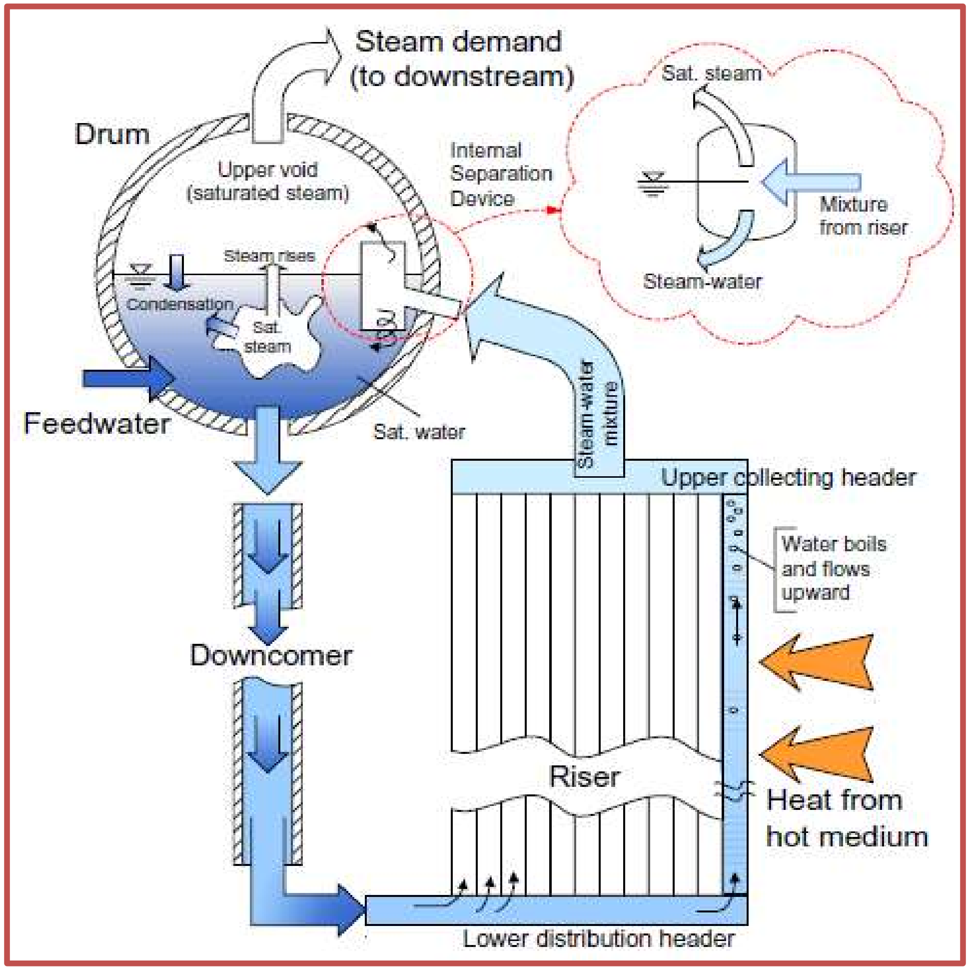


Forced circulation vertical reboiler (Arneth and Stichlmair, 2001)



Two phase flow in vertical thermosyphon reboiler reboiler (Arneth and Stichlmair, 2001)

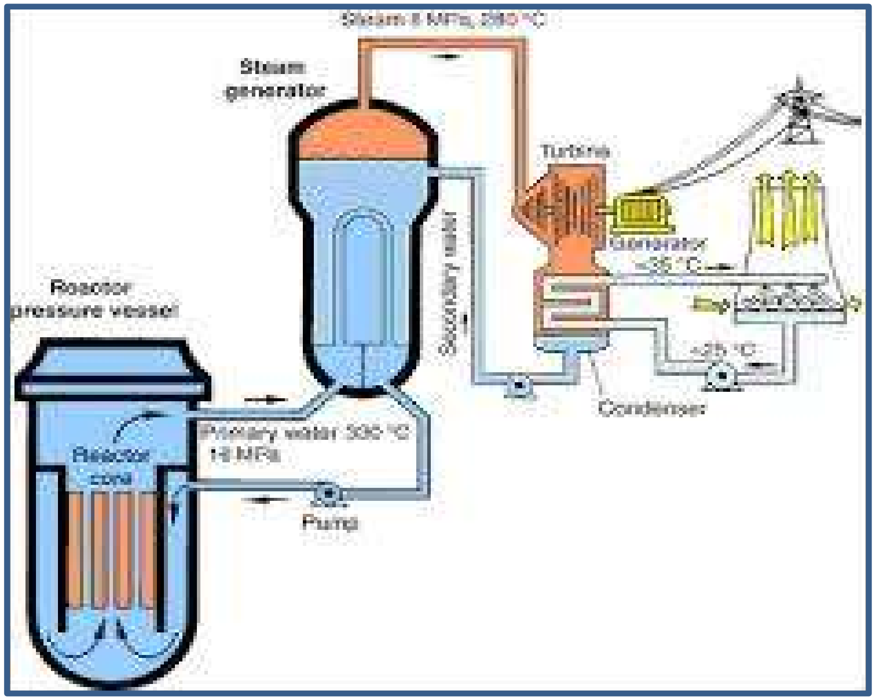


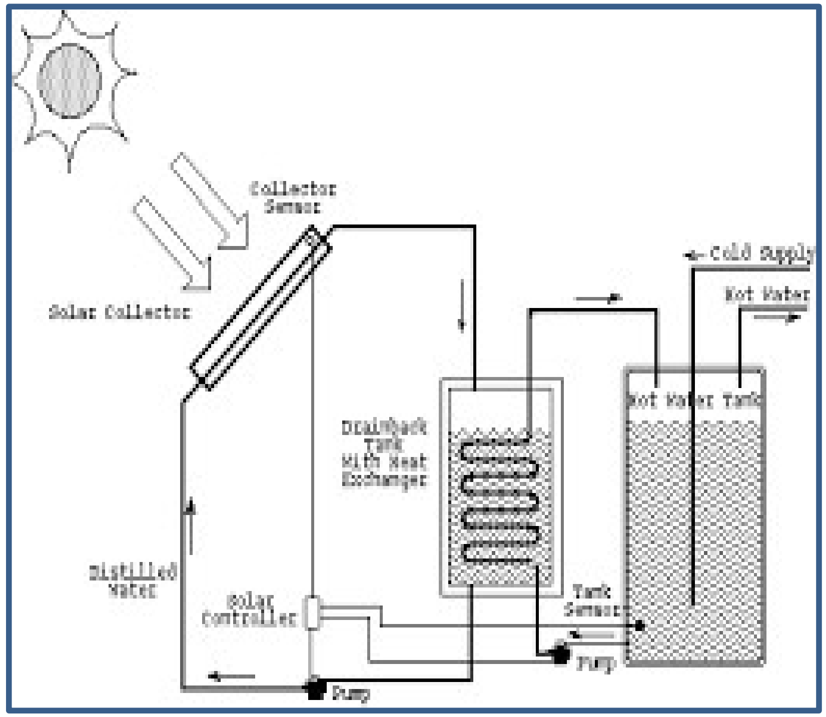


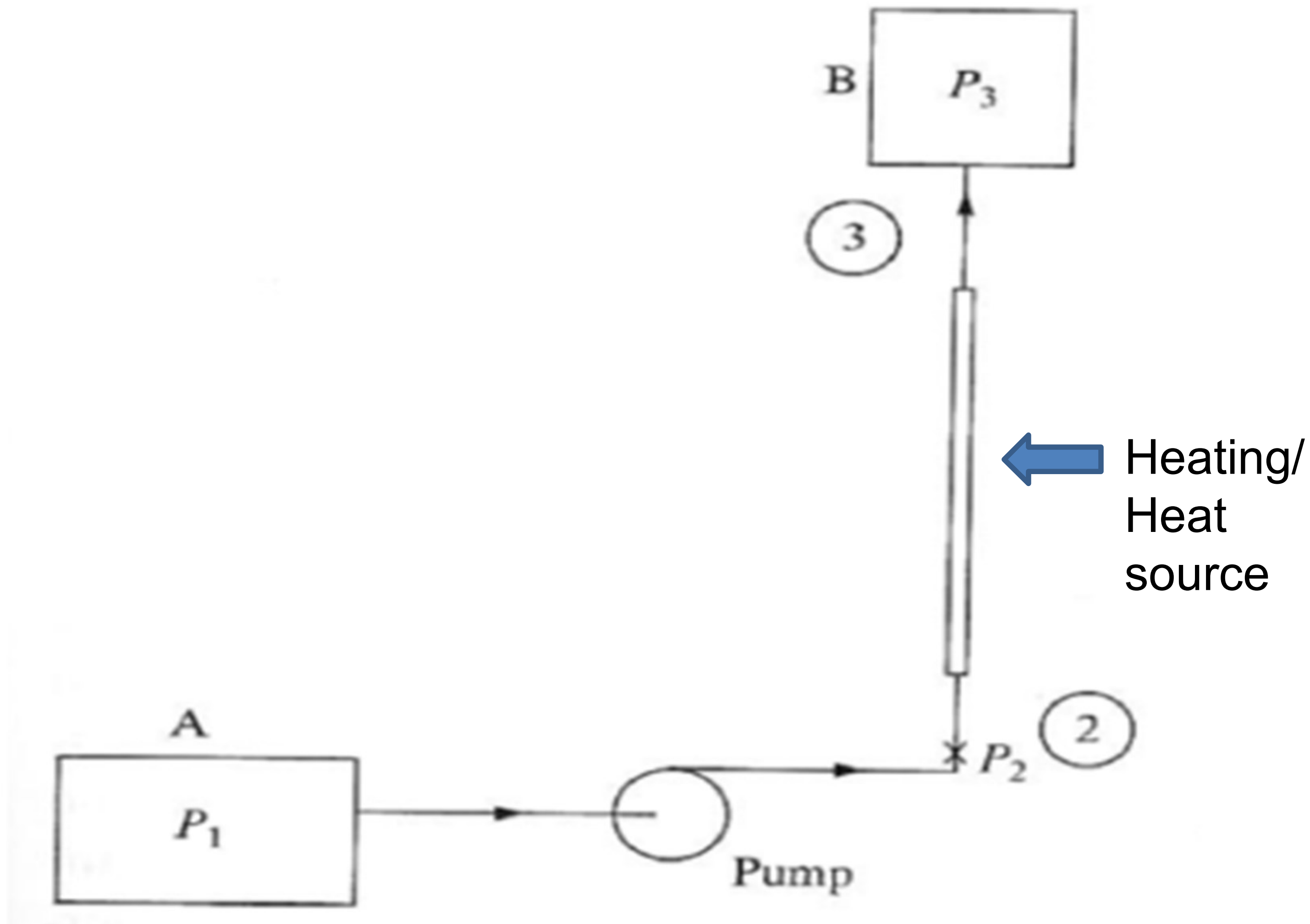
Natural circulation cooled ethylene oxide chemical reactor (Nayak et al., 2006)

Natural circulation drum type boiler (Kim and Choi, 1995)

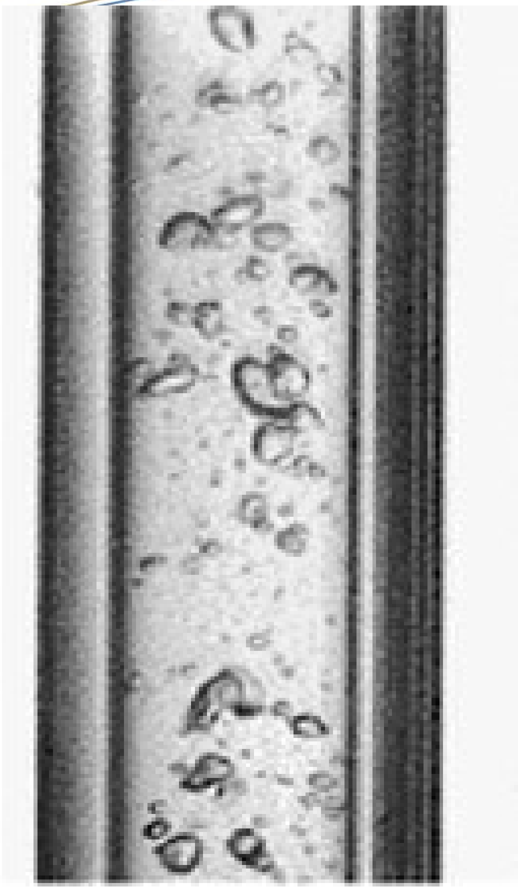
- Nuclear reactor cooling system
- Solar water heater



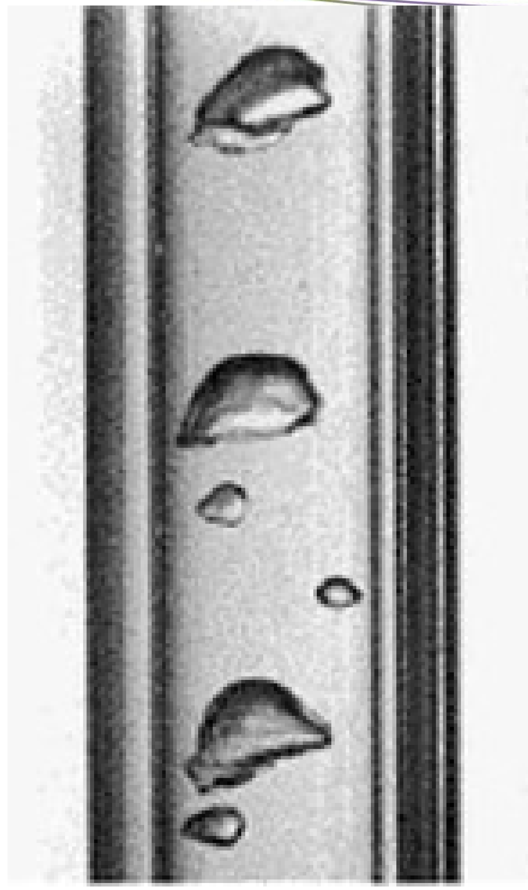




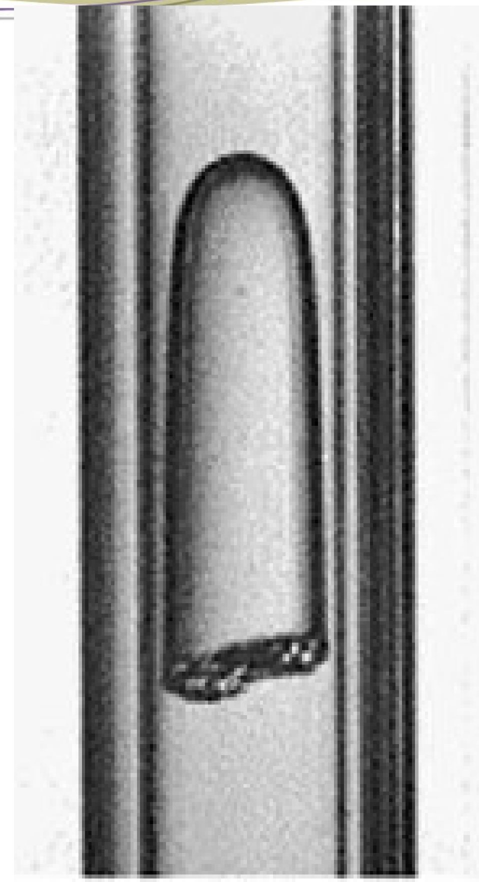
Sketch of boiling system



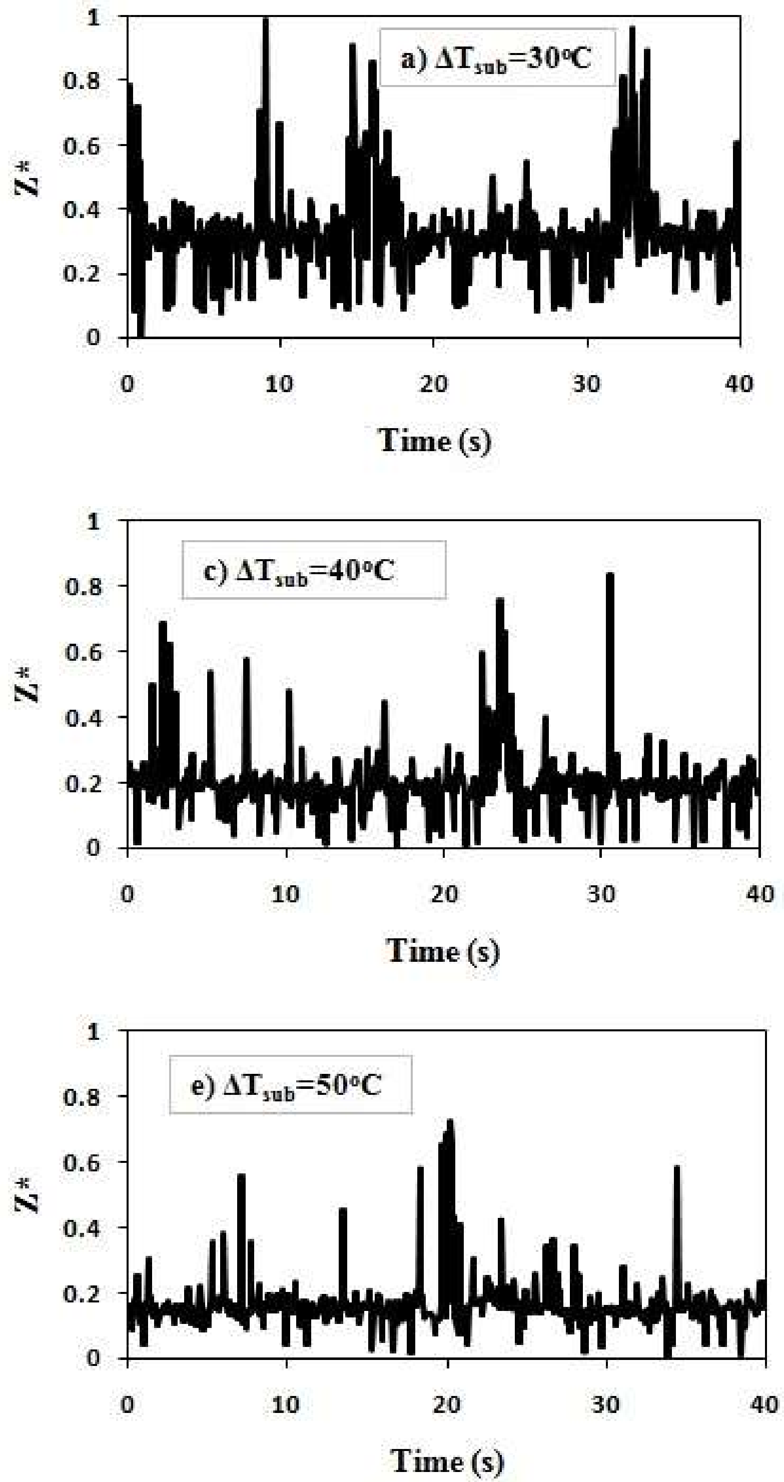
Dispersed Bubble



Bubbly flow



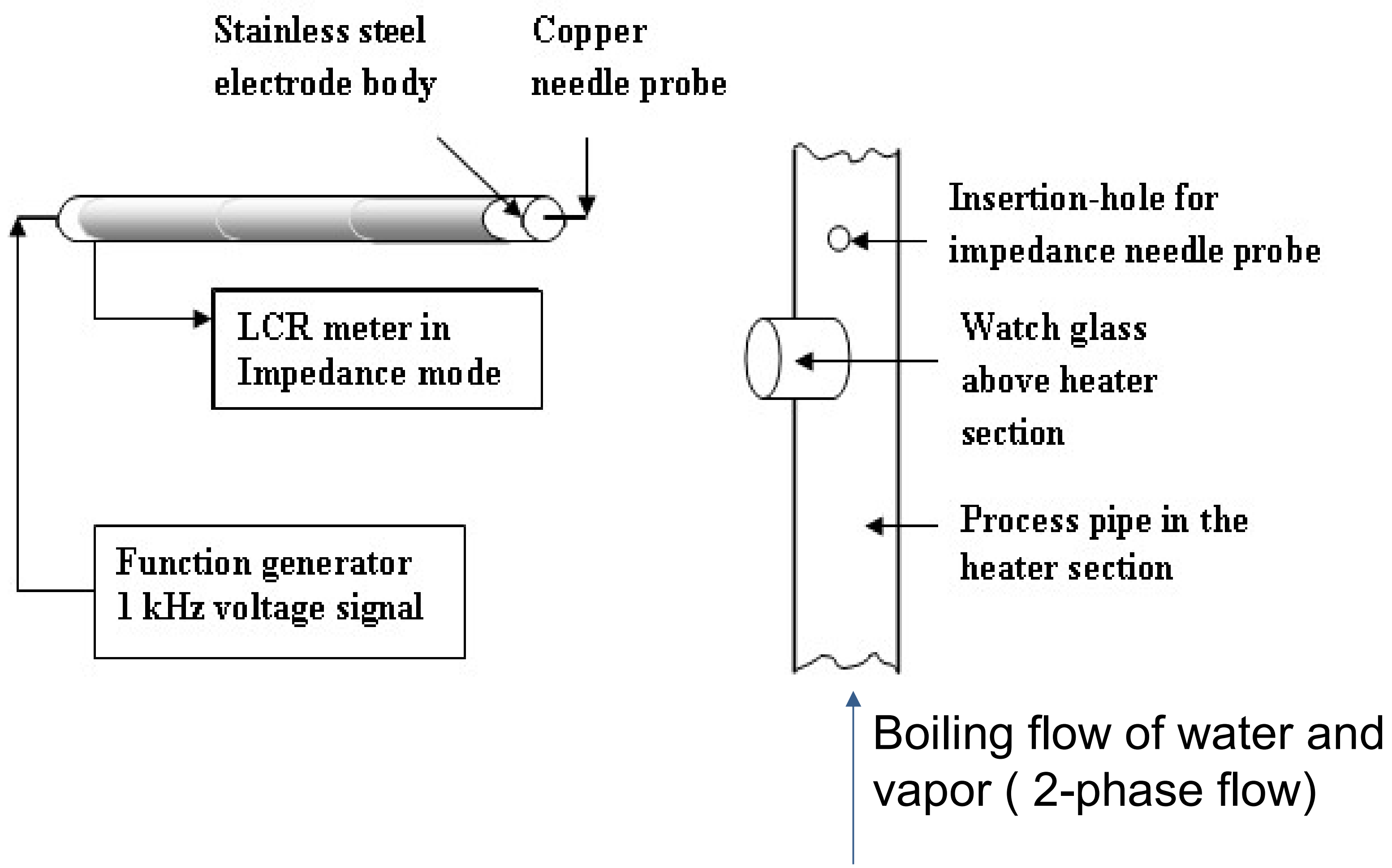
Slug flow



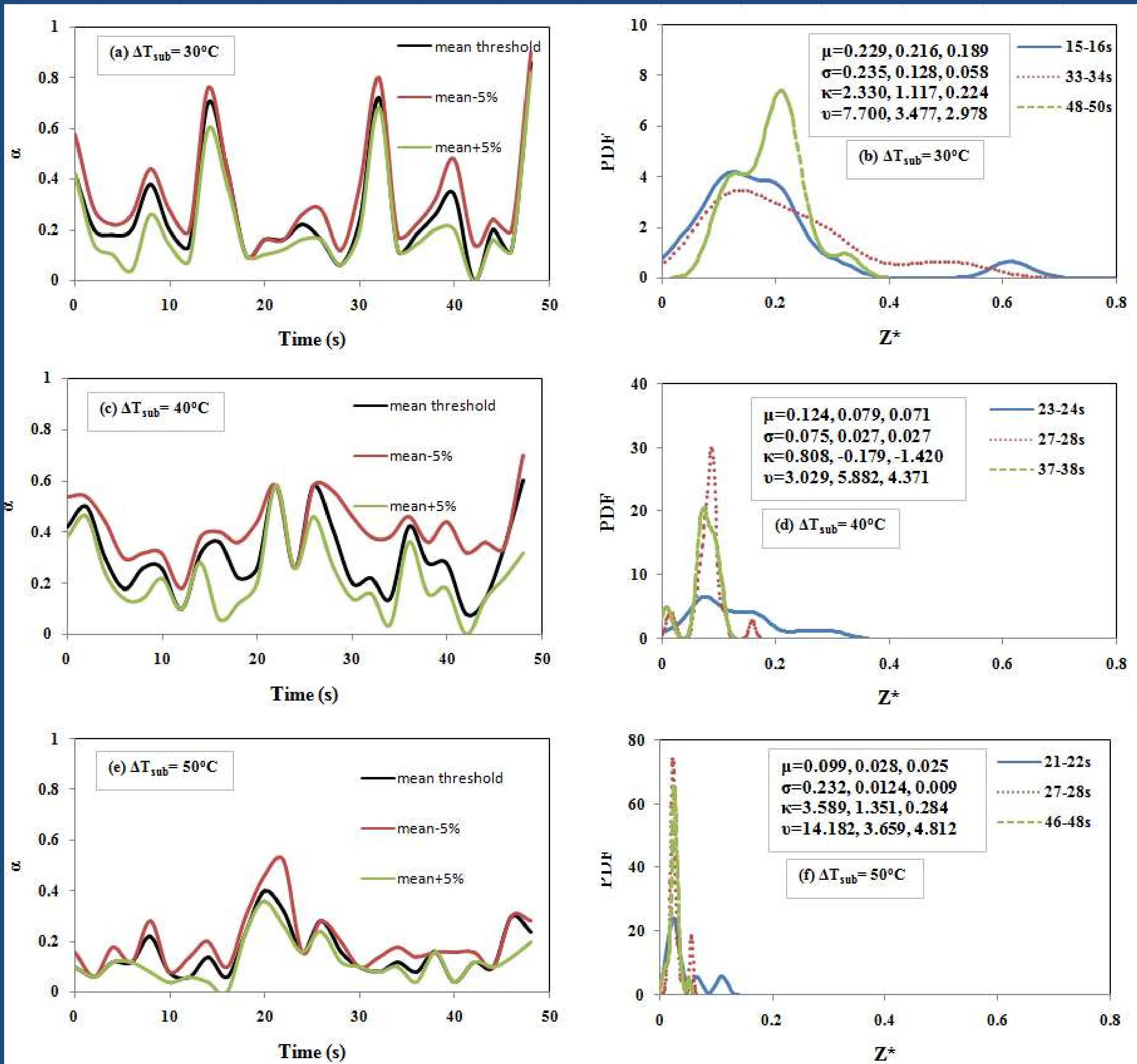
Experimental voltage and impedance signals (Z*) are used for detecting vapor phase and its pattern through the estimation of Probability density function PDF.

PDF of these experimental datapoints is a flow regime identifier that helps to estimate 2-phase heat transfer coefficient.

Experimental method of vapor phase detection in 2-phase flow.



Estimation of void fraction



Slug flow

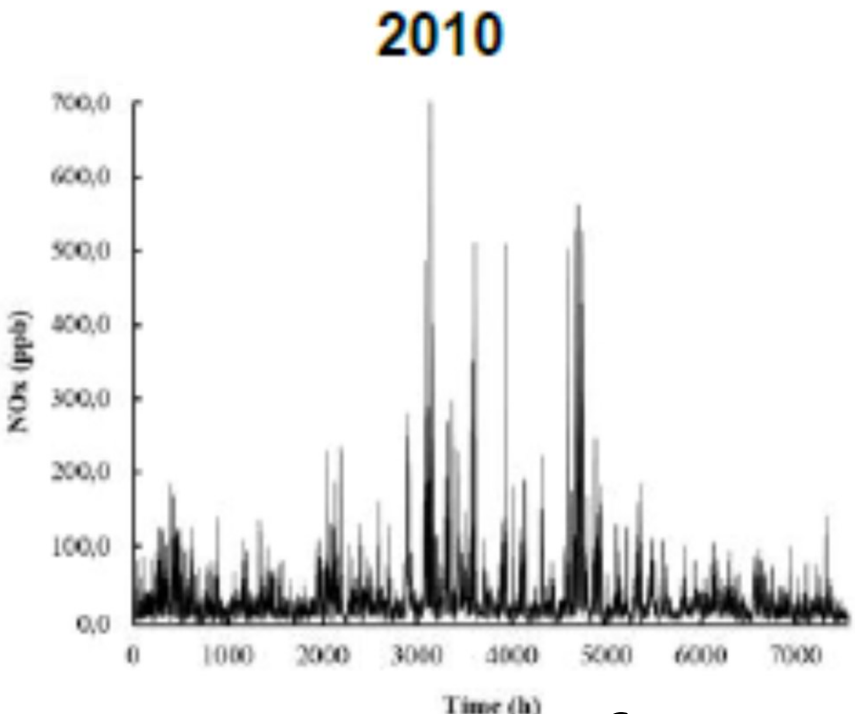
Void fraction (α) estimation

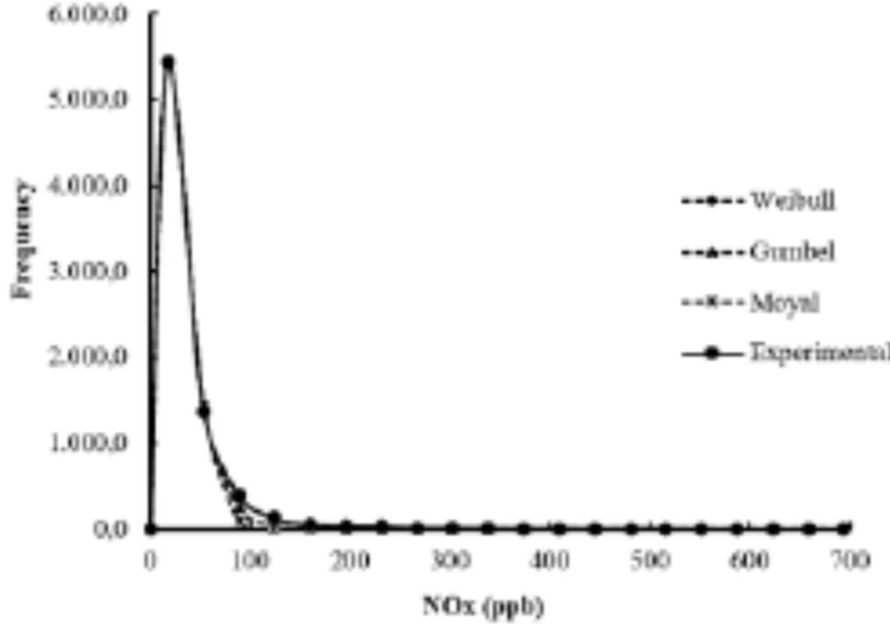
$$\alpha(z,\Delta t) = \frac{1}{\Delta t}\int_0^{\Delta t} M_k(z,t)\,dt$$

Flow regime identification from PDF of impedance signal and void fraction in boiling flow.

Bubbly flow

Oscillations of estimated $\alpha(z,\Delta t)$ with the PDFs of impedance signal at the peaks of the oscillations of void fraction and statistical parameters at different inlet subcoolings.





- *Time series of NOx concentration and its fitted model for the years 2010*

Wesley H. Prieto, Marco A. Cremasco, Application of Probability Density Functions in Modelling, Annual Data of Atmospheric NOx Temporal Concentration, ***CHEMICAL ENGINEERING TRANSACTIONS, VOL. 57, 2017***

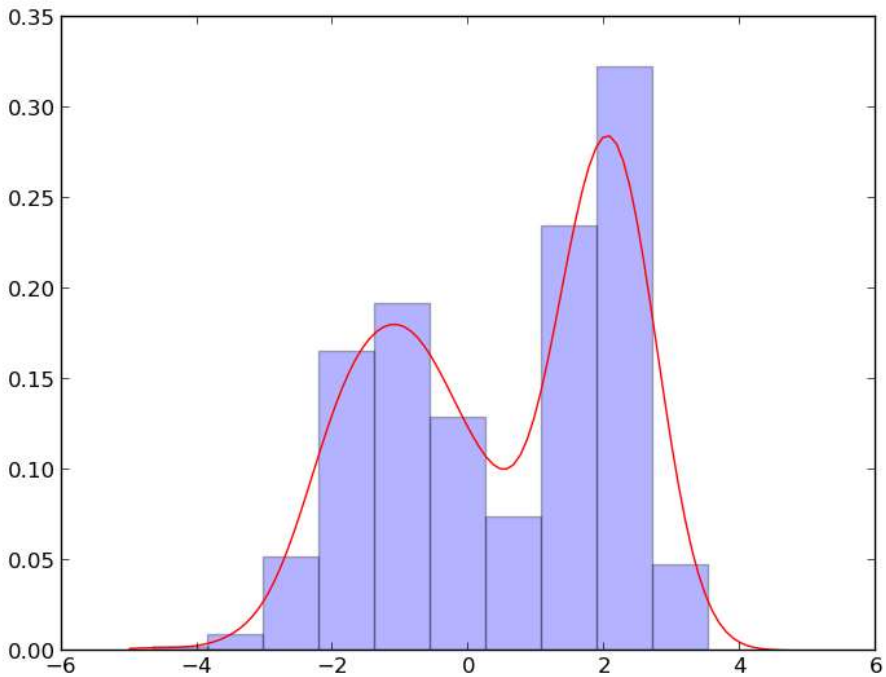
Table 1: Mean, standard deviation, variance, skewness and kurtosis calculated for each time series.

	2010	2011	2012	2013	2014	2015
μ (ppb)	35.90	31.44	26.78	26.72	26.38	22.09
σ (ppb)	47.44	41.17	29.27	31.79	32.28	24.11
σ^2	2250.23	1694.79	856.77	1010.86	1042.15	581.14
A	5.37	4.61	4.03	4.40	5.06	4.54
C	41.14	29.70	25.18	27.14	38.54	30.18

- Wesley H. Prieto, Marco A. Cremasco, Application of Probability Density Functions in Modelling, Annual Data of Atmospheric NOx Temporal Concentration, ***CHEMICAL ENGINEERING TRANSACTIONS, VOL. 57, 2017***

Kernel probability density function of signal

- It is a non parametric method for estimation of probability density function.
- X=[x1 x2 x3…….xn] is n number of random variable.
- PDF is constant in the small interval of dx (piecewise continuous) and n_{dx} is the number of sampling falling in the interval.
- $P_{dx} = \frac{n_{dx}}{n};$



Continuous PDF

Normal: $$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2} \right)$$

Log-Normal: $$f(x) = \frac{1}{x\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(\ln(x)-\mu)^2}{2\sigma^2} \right)$$

Moyal: $$f(x) = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma} + \exp\left(-\frac{x-\mu}{\sigma} \right) \right) \right]$$

Gumbel: $$f(x) = \frac{1}{a} \exp\left[-\frac{x-b}{a} - \exp\left(\frac{x-b}{a} \right) \right]$$

Weibull: $$f(x) = \frac{\eta}{b}\left(\frac{x}{b} \right)^{(\eta-1)} \exp\left[-\left(\frac{x}{b} \right)^{\eta} \right]$$

Gama: $$f(x) = a(ax)^{b-1} \frac{1}{\Gamma(b)} \exp(= ax)$$

Rayleigh: $$f(x) = \frac{1}{a^2} \exp\left[-\frac{1}{2}\left(\frac{x}{a} \right)^2 \right]$$

Maxwell: $$f(x)=\frac{1}{a^3}\left(\frac{2}{\pi}\right)^{\frac{1}{2}}x^2\exp\left[-\frac{1}{2}\left(\frac{x}{a}\right)^2\right]$$

Logistic: $$f(t)=\frac{1}{k}\exp\left(\frac{t-\alpha}{k}\right)\left(1+\exp\left(\frac{t-\alpha}{k}\right)\right)^{-2}$$

Cauchy: $$f(x)=\frac{1}{\pi\left(1+x^2\right)}$$

Pearson III: $$f(x)=\frac{1}{a\Gamma(b)}\left(\frac{x-c}{a}\right)^{b-1}\exp\left[-\left(\frac{x-c}{a}\right)\right]$$

Exponential: $$f(x)=\frac{1}{a}\exp\left[-\left(\frac{x}{a}\right)\right]$$

Chi-Square: $$f(x)=\left(\frac{x}{2}\right)^{\left(\frac{a}{2}-1\right)}\frac{1}{2\Gamma\left(\frac{a}{2}\right)}\exp\left(-\frac{x}{2}\right)$$

Beta: $$f(x)=\frac{1}{\frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}}x^{(a-1)}(1-x)^{(b-1)}$$

Properties of PDF

Mean: $\mu = \dfrac{\sum_{j=1}^{k} f_j x_j}{N}$

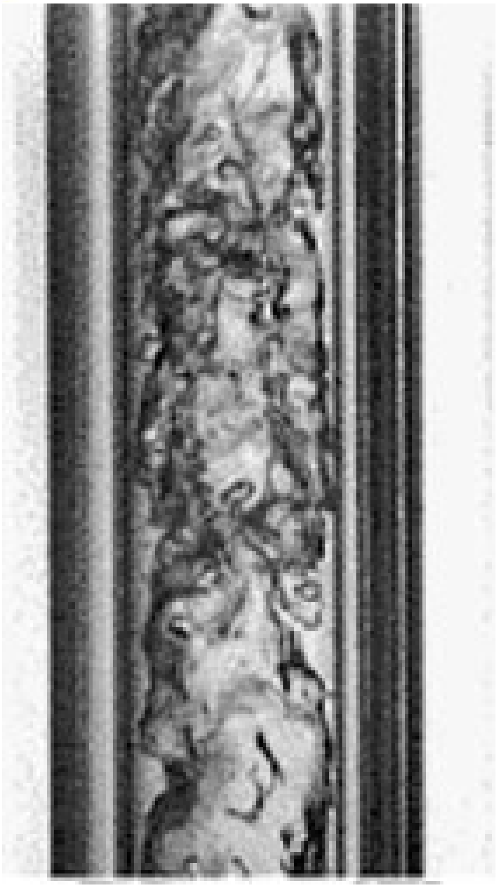
Standard Deviation: $\sigma = \sqrt{\dfrac{\sum_{j=1}^{N} (x_j - \mu)^2}{N}}$

Variance: $\sigma^2 = \sqrt{\dfrac{\sum_{j=1}^{N} (x_j - \mu)^2}{N}}$

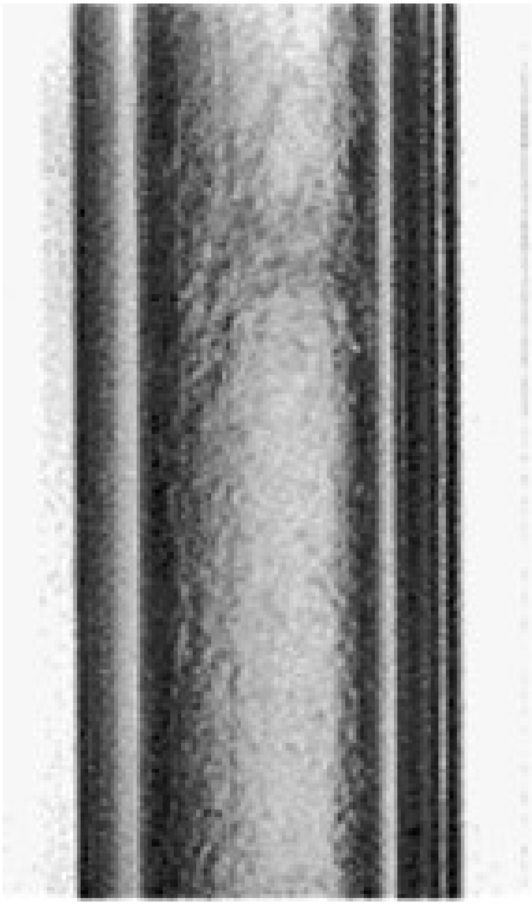
Skewness: $A = \dfrac{E(X-\mu)^3}{\sigma^3}$

Kurtosis: $C = \dfrac{E(X-\mu)^4}{\sigma^4}$

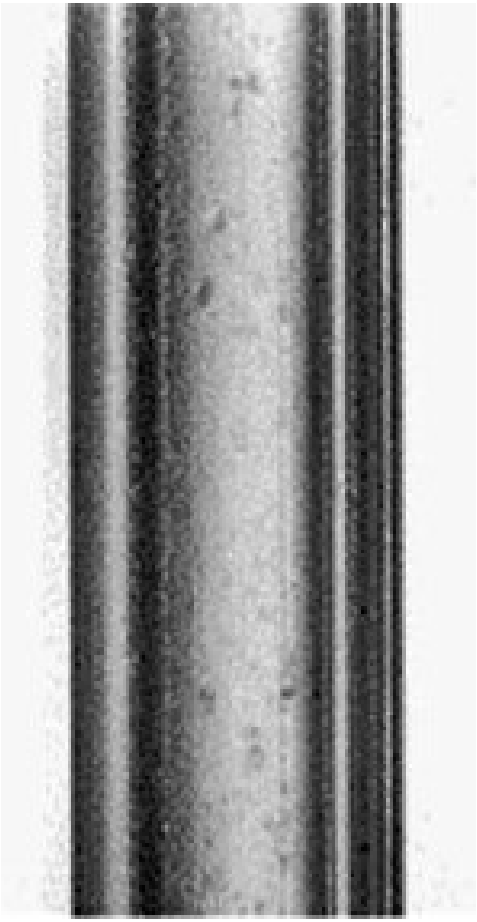
Wesley H. Prieto, Marco A. Cremasco, Application of Probability Density Functions in Modelling Annual Data of Atmospheric NOx Temporal Concentration, ***CHEMICAL ENGINEERING TRANSACTIONS, VOL. 57, 2017***



Churn Flow



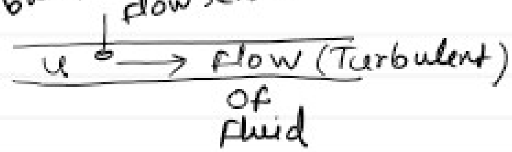
Annular Flow



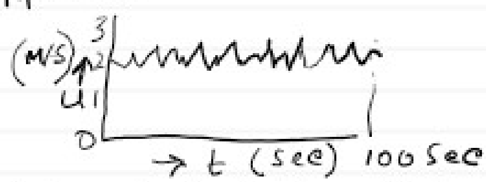
Mist flow

Appendix for continuous PDF estimation (Normal PDF)

$T = 25°C$ } – Fixed.
$P = 1$ bar }



flow signal with time.



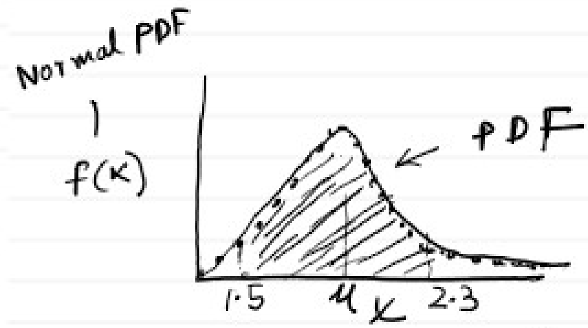
200 data at the interval of 0.5 s.

$\mu = 2$ m/s. (mean value)

$\sigma = 0.2$ m/s (Standard deviation)

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

where x = velocity (change with time)



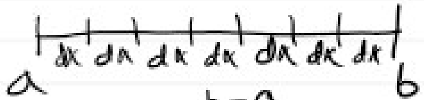
What is the probability of velocity in the range of 1.5 to 2.3

$$= \int_{1.5}^{2.3} f(x)\, dx$$

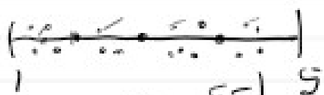
Appendix for experimental PDF estimation (Kernel PDF)

$b = 5$, $a = 1$, $dx = 1$

$\therefore n = \frac{5-1}{1} = 4$



$n+1 = \frac{b-a}{dx}$ $dx = \frac{b-a}{n+1}$



n = no. of segments.

$\Delta x = \frac{5-1}{3+1} = (1)$

$\Delta x = \frac{b-a}{n+1}$

$P_1 = \frac{4}{n_{Total}}$; $P_2 = \frac{3}{n_{total}}$; $P_3 = \frac{5}{n_{total}}$.

$P_4 = \frac{4}{n_{total}}$.

n_{total} = Total no. of data point.

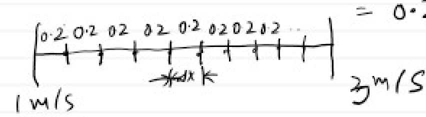
$n_{total} = 200$

in the range of 1 m/s ... 3 m/s.

n = 9 subdivision

$\Delta x = \frac{3-1}{9+1} = \frac{2}{10} = \frac{1}{5} = 0.2$

interval



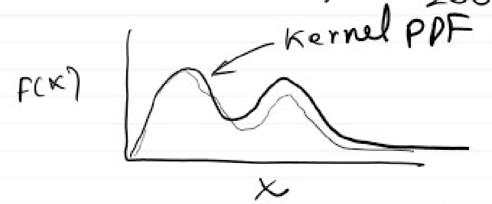
(1 – 1.2) $P_1 = \frac{3}{200}$,

(1.2 – 1.4) $P_2 = \frac{5}{200}$

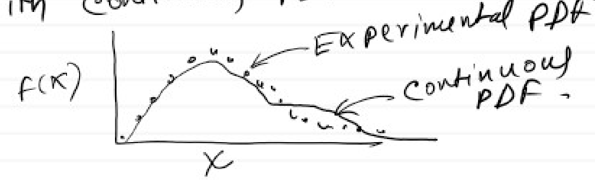
(1.4 – 1.6) $P_3 = \frac{10}{200}$

$P_4 = \frac{20}{200}$

$P_9 = \frac{n_9}{200}$



One has to match or fit experimental PDF with continuous PDF.



Thank you