25~26 Feb. 2005 Cherry Bud Workshop 2005（Keio University）

An approach to the extreme value distribution of non-stationary process

Hang CHOI & Jun KANDA

Institute of Environmental Studies

Graduate School of Frontier Sciences

The University of Tokyo

01 Theoretical Frameworks of EVA

1) Stationary Random Process (Sequence)

- Distribution Ergodicity
- (In)dependent and Identically Distributed random variables (I.I.D. assumption)

→ strict stationarity

* Conventional approach

2) Non-stationary Random Process (Sequence)

- (In)dependent but non-identically Distributed random variables (non-I.I.D. assumption)

→ weak stationarity, non-stationarity

3) Ultimate (Asymptotic) and penultimate forms

1) I.I.D. random variable Approach

Ultimate form : Fisher-Tippett theorem (1928)

$$\underline{X_1,\mathrm{K}\ ,X_n \in F(x)},\ Z_n = \max\left\{X_1,\ldots,X_n\right\}$$

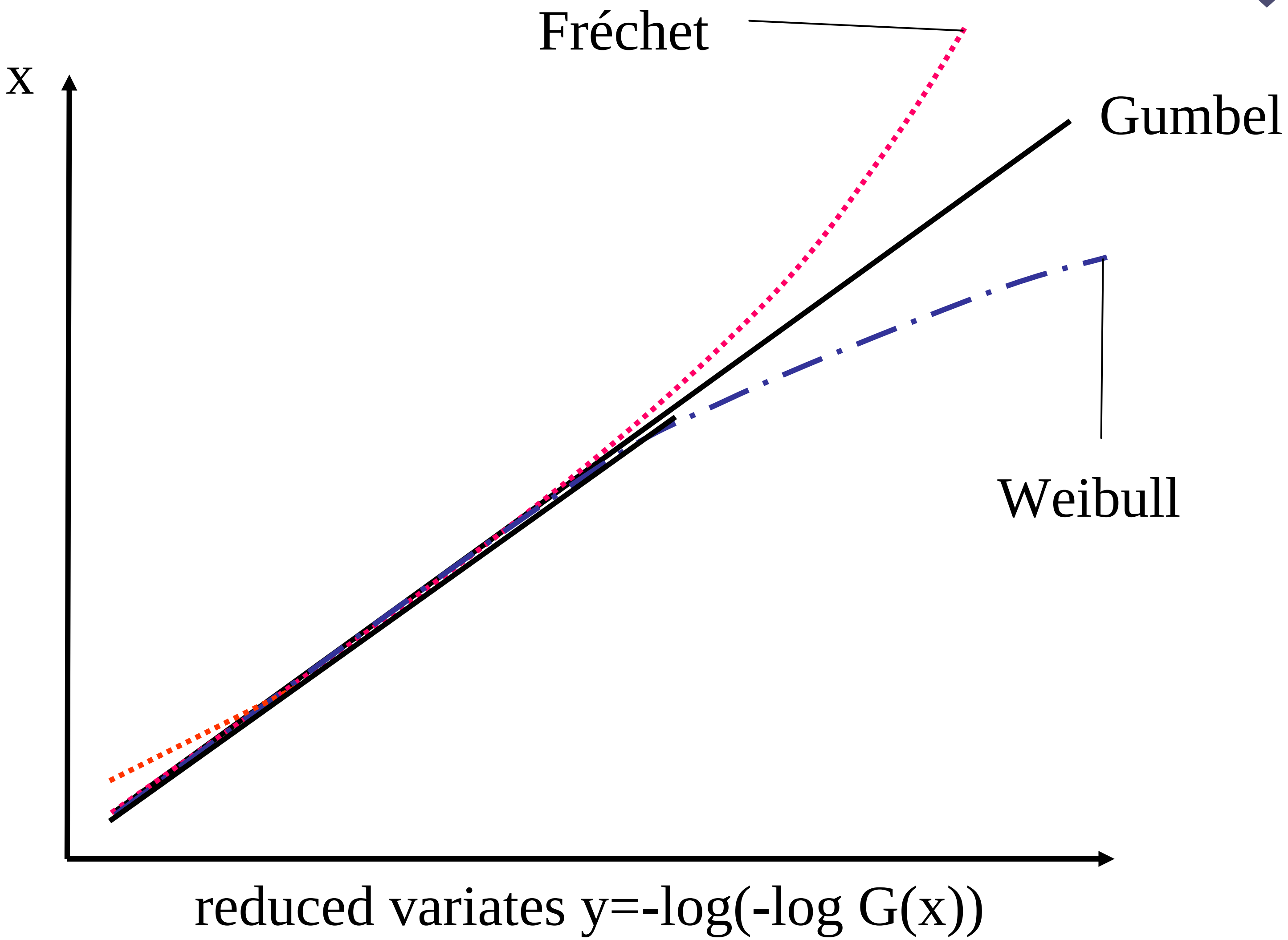
$$\lim_{n\to+\infty} P\left(\frac{Z_n - a_n}{b_n} \le x\right) = \lim_{n\to+\infty} \underline{F^n(a_n + b_n x) = G(x)} \in \mathrm{F}$$

$$\mathrm{F} = \{\text{ Gumbel, Fréchet, Weibull }\}$$

↓

GEVD, GPD, POT-GPD + MLM, PWM, MOM etc.

R.A. Fisher & L.H.C. Tippett (1928), Limiting forms of the frequency distribution of the largest or smallest member of a sample, *Proc. Cambridge Philosophical Society*, Vol 24, p180~190


Fréchet
x
Gumbel
Weibull
reduced variates y=-log(-log G(x))

2) non-I.I.D. random variable Approach

$$\underline{X_1 \in F_1(x), \ldots, X_n \in F_n(x),}\ Z_n = \max\{X_1, \ldots, X_n\}$$

$$\lim_{n \to +\infty} P\left(\frac{Z_n - c_n}{d_n} \leq x\right) = \lim_{n \to +\infty} \underline{\prod_{j=1}^{n} F_j(d_j + c_j x)} = Q(x) \in \mathrm{S}$$

$$\mathrm{F} \subset \mathrm{S}$$

$\mathrm{S} =$ { Gumbel, Fréchet, Weibull,…..}

The class of EVD for non-i.i.d. case is much larger.

* Falk *et al.*, *Laws of Small Numbers: Extremes and Rare Events*, Birkhäuser, 1994

3) Ultimate / penultimate form and finite epoch T in engineering practice

In engineering practice, the epoch of interest, $\boldsymbol{T}$ is finite.
e.g. annual maximum value, monthly maximum value and maximum/minimum pressure coefficients in 10min etc.

As such, the number of independent random variables m in the epoch T becomes a finite integer, *i.e.* $m<\infty$, and consequently, the theoretical framework for ultimate form is no longer available regardless *i.i.d.* or *non-i.i.d.*case.

Following discussions are restricted on the *penultimate d.f.* for the extremes of non-stationary random process *in a finite epoch T*.

02 Statement of the Problem

Let $x(t),\ t \in \circ^{+} := [0,\infty)$ be the continuous observation record of a non-stationary continuous stochastic process $X(t)$ and assume that $X(t)$ is a mean square differentiable process and hold the following condition.

$$r(t,\tau) := E\left[X(t)X(t+\tau)\right] \Rightarrow r(t,\tau)\log\tau \to 0 \text{ as } \tau \to T$$

As such, how to estimate the extreme value distribution of $X(t)$ in an epoch $T< \infty$ from the continuous observation record $x(t)$?

03 Assumptions and Formulation

According to the conventional approach in wind engineering, let assume the non-stationary process $X(t)$ can be partitioned with a finite epoch T, in which the partitioned process $X_i(t), (i-1)T \leqq t < iT$, can be assumed as an independent stationary random process, and define the d.f. $F_Z(x)$ as follows.

$$F_Z(x) = P(Z \le x;\ Z := \sup X(t),\ t \in [0, T < \infty))$$

Then, by the Glivenko-Cantelli theorem and the block maxima approach

$$F_Z(x) = \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n} I_{(-\infty,x]}\big(Z_i := \sup X_i(t)\big) = \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n} P(Z_i \le x)$$

where $I_c(x) = 1$ if $x \in c$ else 0

04 i.i.d. random sequence (EQRS) approach of $P(Z_i \le x)$

By partitioning the interval [(*i*-1)*T*,*iT*) into finite subpartitions $[(j-1)h, jh), 1 \le j \le [T/h] = m_i$ in the manner of that

$$P(Z_i \le x) = \prod_{j=1}^{m_i<\infty} P\left(Z_j^* \le x;\ Z_j^* := \sup\left(X_i(t), t \in [(j-1)h, jh)\right)\right) = F_{Z_i}^{m_i}(x)$$

the required d.f. $F_Z(x)$ can be defined as follows.

$$F_Z(x) := \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} F_{Z_i}^{m_i}(x)$$

If it is possible to assume that all $m_i \approx m$, then

$$F_Z(x) := \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} F_{Z_i}^{m}(x)$$

05 lower bound of $F_Z(x)$

From the inequality (geometric mean)<(arithmetic mean), a lower bound of $F_Z(x)$ can be defined as follows:

$$\tilde{F}_Z(x) := \lim_{n\to\infty} \prod_{i=1}^{n} F_{Z_i}^{m/n}(x) < \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} F_{Z_i}^{m}(x) = F_Z(x)$$

06 alternative definition of $F_Z(x)$

$$\hat{F}_Z(x) := \left(\frac{1}{n} \sum_{i=1}^{n} F_{Z_i}(x) \right)^{m} = \bar{F}_Z^{m}(x)$$

This definition can be found easily in engineering applications and may be reasonable for the case of $F_{Z_1}\ ;\ \mathrm{L}\ ;\ F_{Z_n}$.

07 Inequality of quantile functions

Let define quantile functions of each definition as follows.

$$Q(\alpha) := F_Z^{-1}(\alpha),\ \tilde{Q}(\alpha) := \tilde{F}_Z^{-1}(\alpha),\ \hat{Q}(\alpha) := \hat{F}_Z^{-1}(\alpha)$$

Then, by the inequality for the means and the comparison of the distribution of order statistics, i.e. $Z_{\mathrm{n:n}}$ and $Z_{\mathrm{1:n}}$,

$$Q(\alpha) \le \tilde{Q}(\alpha),\ \hat{Q}(\alpha) < \tilde{Q}(\alpha) \quad \mathit{for\ all}\ \alpha \in (0,1)$$

$$\begin{cases} \hat{Q}(\alpha) < Q(\alpha) \to \tilde{Q}(\alpha) & \text{for large } \alpha \\ Q(\alpha) < \hat{Q}(\alpha) < \tilde{Q}(\alpha) & \text{for small } \alpha \end{cases}$$

Therefore, the alternative definition results in smaller variance of extremes.

Complement for the inequalities of quantile functions

① $$Z_{n:n} \sim \prod_{i=1}^{n} F_i^{m},\ \tilde{Z}_{n:n} \sim \left(\prod_{i=1}^{n} F_i^{m/n}\right)^{n} = \prod_{i=1}^{n} F_i^{m} \Rightarrow Z_{n:n}(\alpha) = \tilde{Z}_{n:n}(\alpha)$$

② $$\tilde{Z}_{n:n} \sim \prod_{i=1}^{n} F_i^{m},\ \bar{Z}_{n:n} \sim \left(\frac{1}{n}\sum_{i=1}^{n} F_i\right)^{mn},\ \left(\prod_{i=1}^{n} F_i^{m/n}\right)^{n} < \left(\frac{1}{n}\sum_{i=1}^{n} F_i\right)^{mn} \Rightarrow \tilde{Z}_{n:n}(\alpha) > \bar{Z}_{n:n}(\alpha)$$

③ $$Z_{1:n} \sim 1-\prod_{i=1}^{n}\left(1-F_i^{m}\right);\ \sum_{i=1}^{n} F_i^{m} - o(F^{2m}),\ \bar{Z}_{1:n} \sim 1-\left(1-\bar{F}^{m}\right)^{n};\ \frac{1}{n^{m-1}}\sum_{i=1}^{n} F_i^{m} - o(F^{2m})$$

$$\Rightarrow Z_{1:n}(\alpha) < \bar{Z}_{1:n}(\alpha)$$

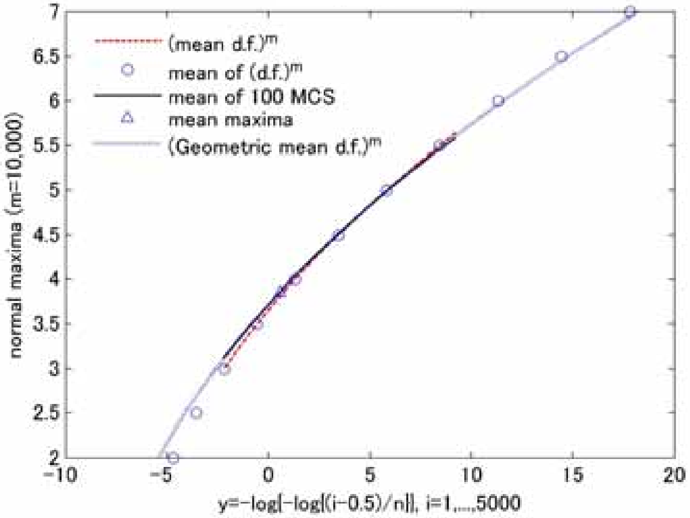
④ $$Z_{1:n} \sim 1-\prod_{i=1}^{n}(1-F_i^{m}),\ \tilde{Z}_{1:n} \sim 1-\left(1-\prod_{i=1}^{n} F_i^{m/n}\right)^{n},$$

$$\prod_{i=1}^{n}(1-F_i^{m}) < \left(1-\bar{F}^{m}\right)^{n} < \left(1-\prod_{i=1}^{n} F_i^{m/n}\right)^{n} \Rightarrow Z_{1:n}(\alpha) < \tilde{Z}_{1:n}(\alpha)$$

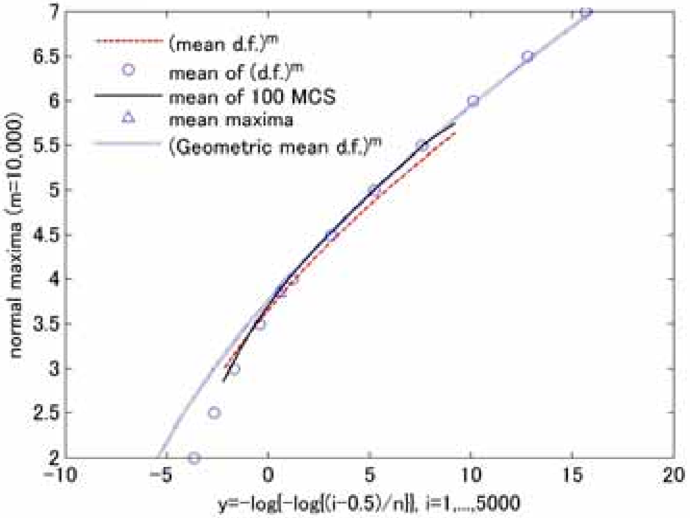
08 Numerical example $\bar{F}(x) = N(0,1)$

(m=10,000, n=5,000, iteration=100)

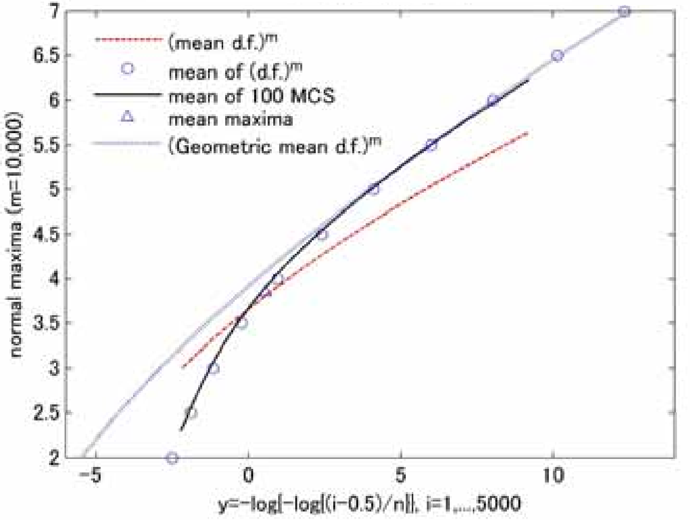
$X_i \sim N(0,1)$



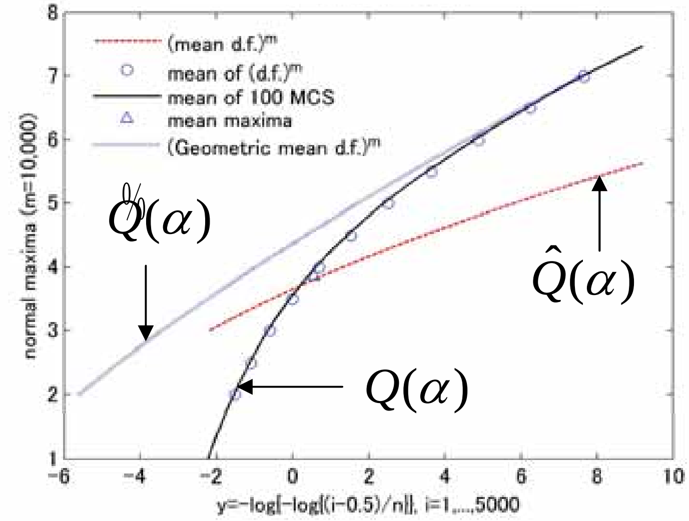
$X_i \sim N(0,\sigma_i),\ \sigma_i \sim N(1,0.05)$



$X_i \sim N(0,\sigma_i),\ \sigma_i \sim N(1,0.1)$



$X_i \sim N(0,\sigma_i),\ \sigma_i \sim N(1,0.2)$



09 practical application:
Annual maximum wind speed in Japan

1) Observation records and Historical annual maximum wind speeds at 155 sites in Japan

Observation Records: JMA records (CSV format)

1961~1990 : 10 min average wind speed per 3 hours

1991~2002 : 10 min average wind speed per hour

Historical annual maximum wind speeds record:

1929~1999 : A historical annual max. wind speed data set compiled by Ishihara *et al*.(2002)

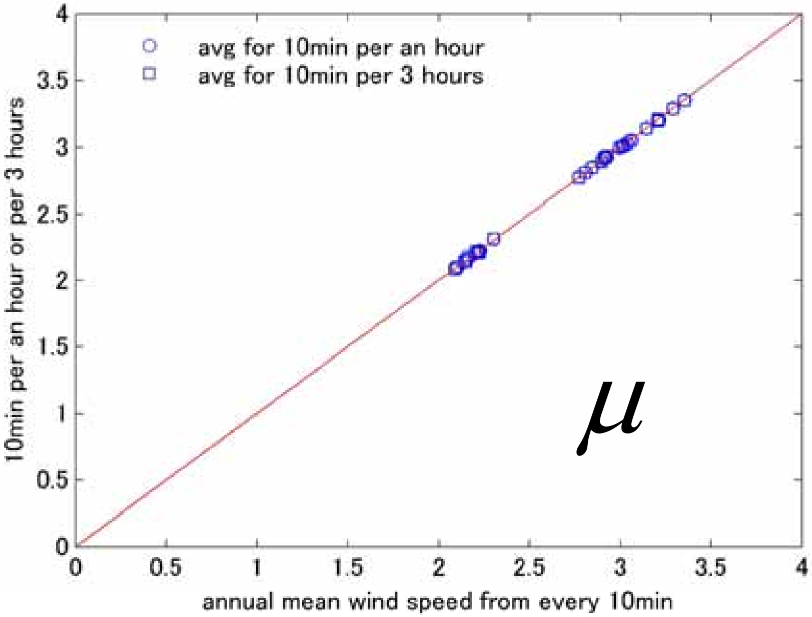
2000~2002 : extracted from JMA records (CSV format)

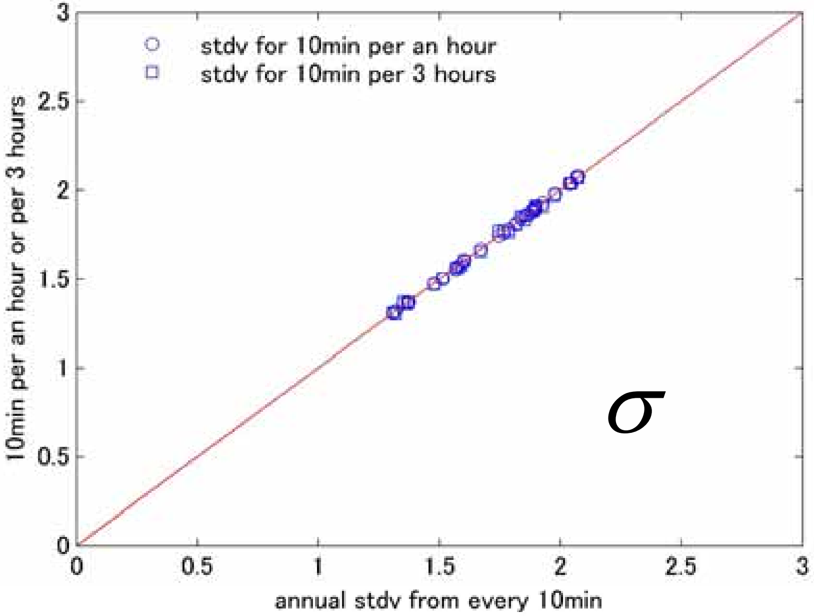
T. Ishihara *et al*. (2002), A database of annual maximum wind speed and corrections for anemometers in Japan, *Wind Engineers*, *JAWE*, No.92, p5~54 (in Japanese)

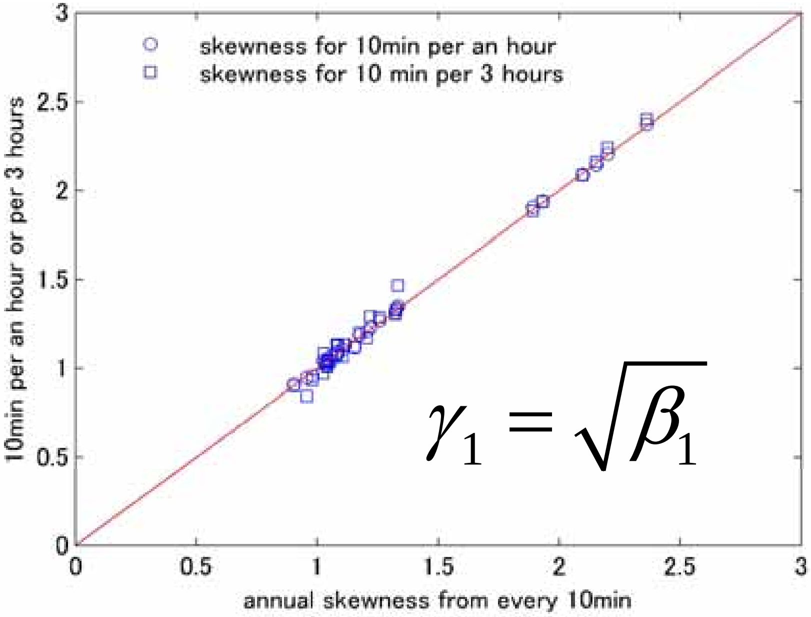
09 practical application:
Annual maximum wind speed in Japan

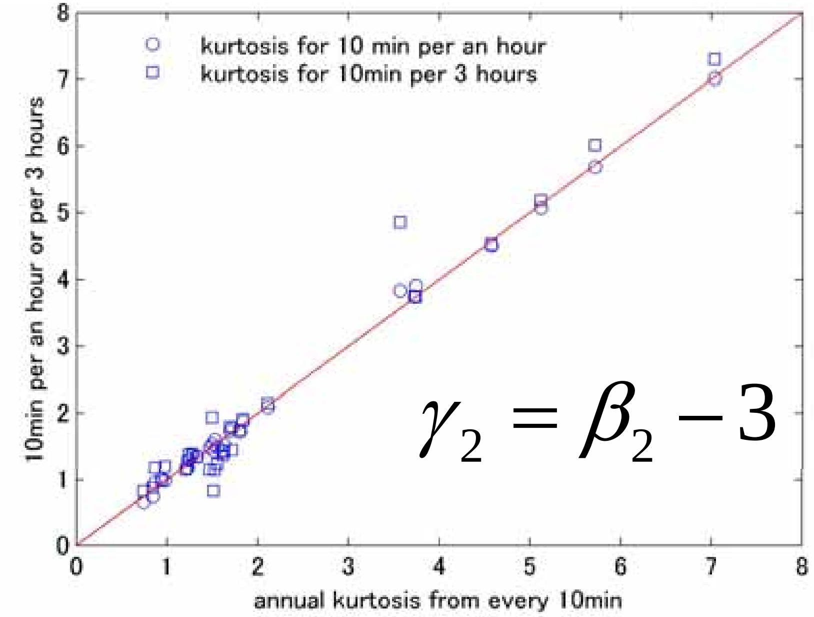
2) The effect of different observation recording format on the basic statistics

Base on the recently opened continuous 1 min average wind speed records (1997. 3~2002. 2), calculating every 10 min average wind speeds, 10 min average per hour and 3 hours, and comparing the basic statistics for each recording format, then the effect of different recording format becomes to be clear.


○ avg for 10min per an hour
□ avg for 10min per 3 hours
μ
10min per an hour or per 3 hours
annual mean wind speed from every 10min
0
0.5
1
1.5
2
2.5
3
3.5
4

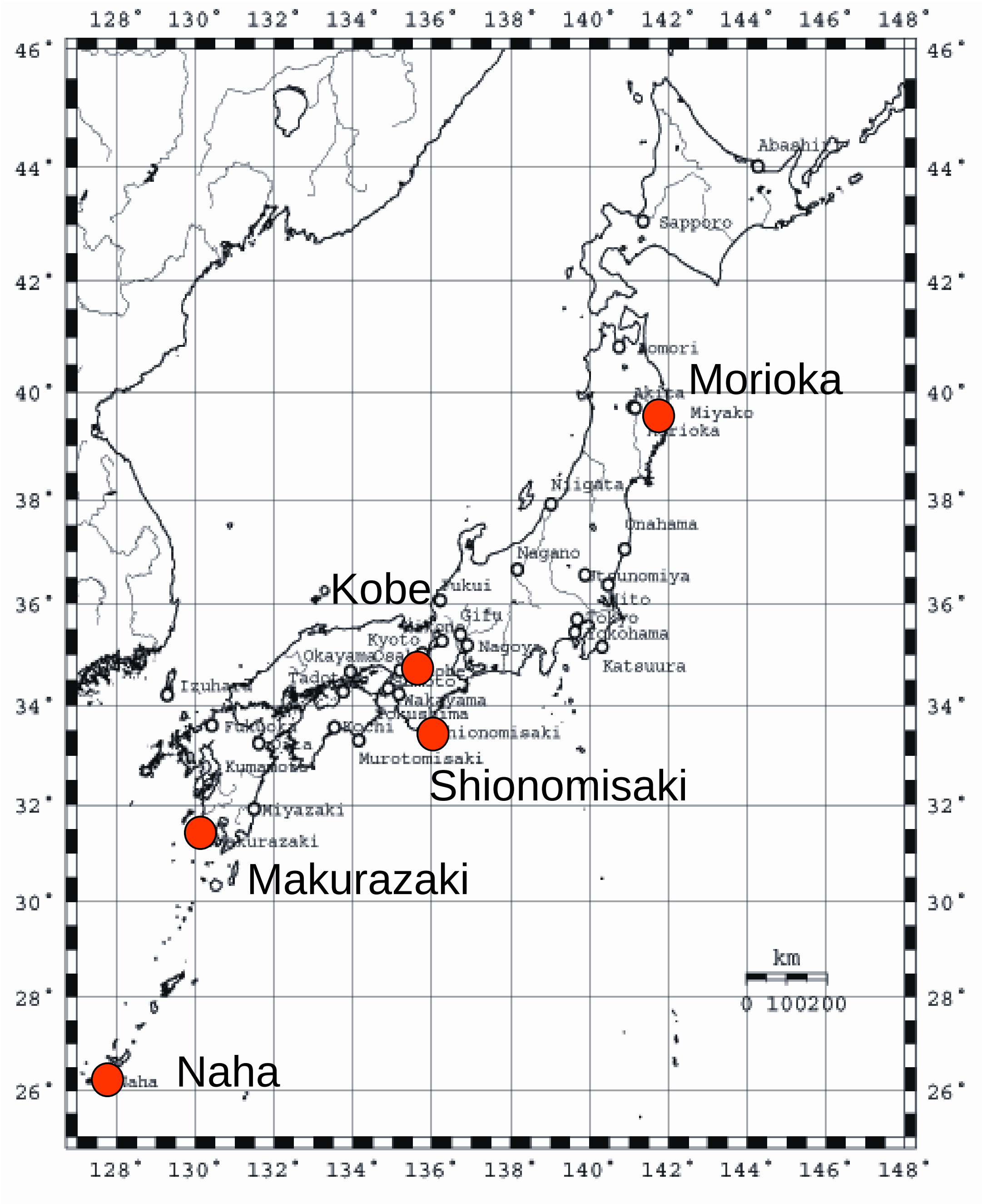

○ stdv for 10min per an hour
□ stdv for 10min per 3 hours
σ
10min per an hour or per 3 hours
annual stdv from every 10min
0
0.5
1
1.5
2
2.5
3


○ skewness for 10min per an hour
□ skewness for 10 min per 3 hours
$\gamma_1 = \sqrt{\beta_1}$
10min per an hour or per 3 hours
annual skewness from every 10min
0
0.5
1
1.5
2
2.5
3

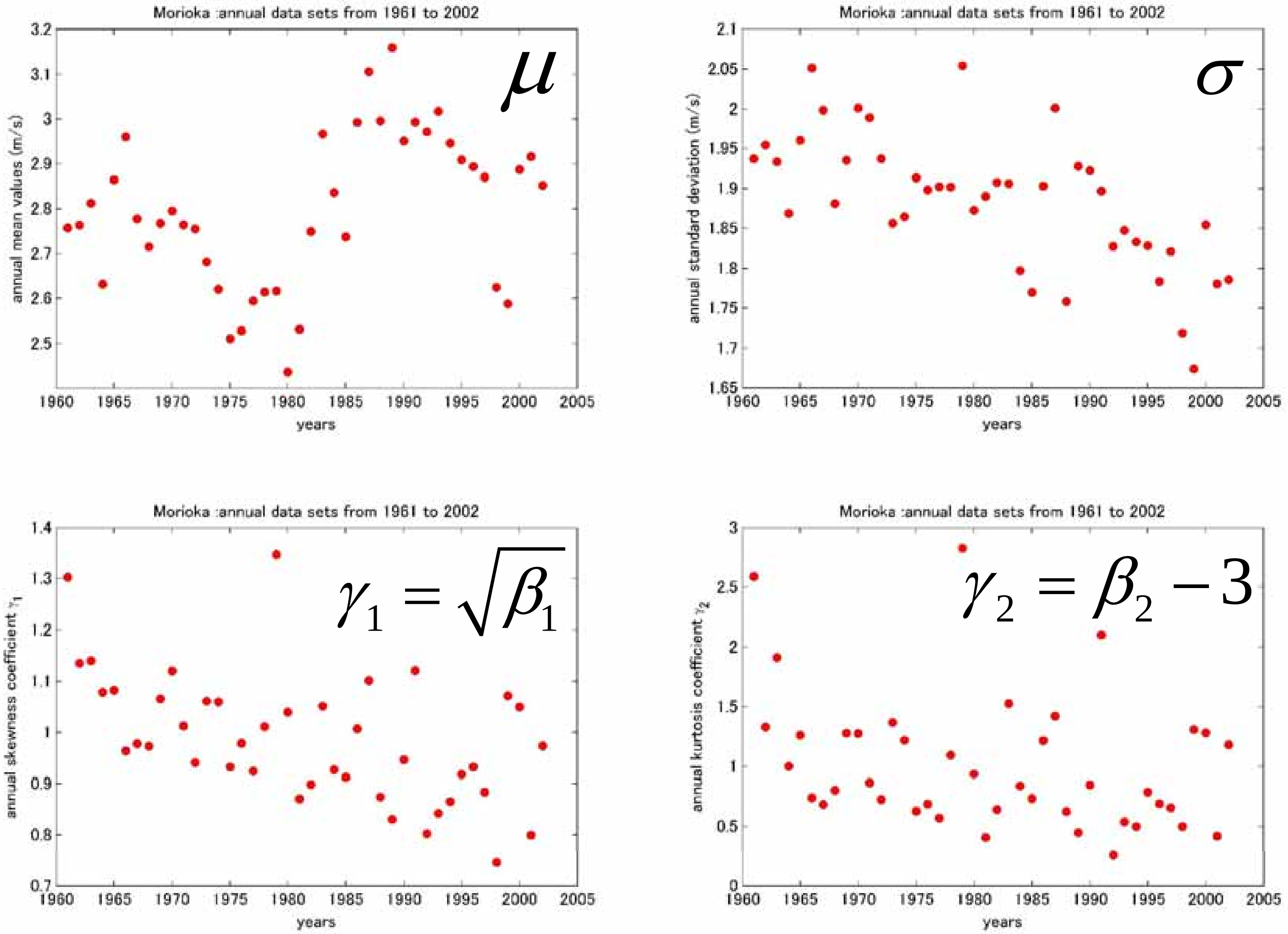

○ kurtosis for 10 min per an hour
□ kurtosis for 10min per 3 hours
$\gamma_2 = \beta_2 - 3$
10min per an hour or per 3 hours
annual kurtosis from every 10min
0
1
2
3
4
5
6
7
8

3) Examples of non-stationarity : the basic statistics (1961~2002)

Site	**Coefficient of Variations (C.O.V)**			
	μ	σ	γ_1	γ_2
Abashiri	5.1%	5.0%	11.7%	14.4%
Katsuura	5.81%	8.87%	24.06%	26.71%
Kobe	5.71%	8.25%	12.55%	16.31%
Kumamoto	7.62%	5.93%	14.19%	32.36%
Makurazaki	4.60%	6.07%	25.97%	47.98%
Morioka	6.99%	4.43%	12.09%	11.55%
Oita	3.42%	8.41%	16.49%	27.41%
Shionomisaki	4.34%	4.07%	19.36%	30.10%
Tokyo	5.90%	8.06%	18.90%	17.88%
Minimum	**3.42%**	**4.07%**	**11.72%**	**11.55%**
Maximum	**7.62%**	**8.87%**	**25.97%**	**47.98%**


Morioka
Kobe
Shionomisaki
Makurazaki
Naha
Abashiri
Sapporo
Aomori
Akita
Miyako
Morioka
Niigata
Onahama
Nagano
Utsunomiya
Fukui
Mito
Gifu
Tokyo
Yokohama
Nagoya
Kyoto
Katsuura
Okayama
Osaka
Kobe
Tadotsu
Izuhara
Sumoto
Wakayama
Tokushima
Fukuoka
Kochi
Shionomisaki
Oita
Murotomisaki
Kumamoto
Miyazaki
Makurazaki
Naha
km
0 100200
128° 130° 132° 134° 136° 138° 140° 142° 144° 146° 148°
46° 44° 42° 40° 38° 36° 34° 32° 30° 28° 26°

Morioka (1961~2002)



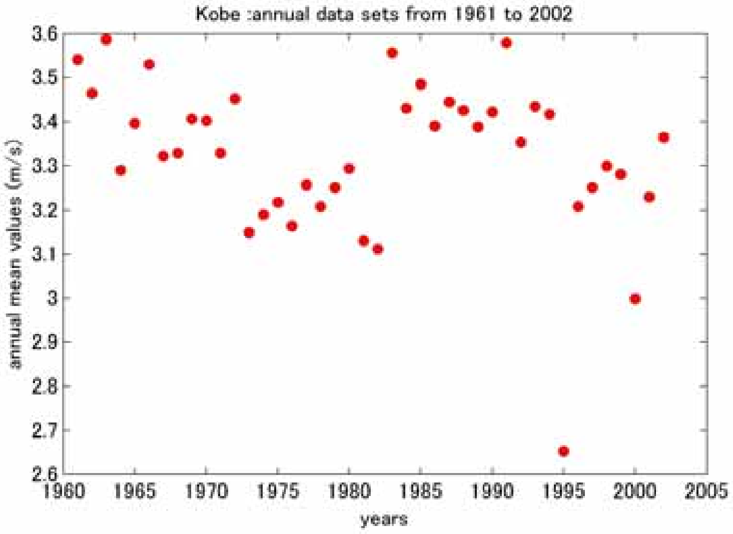
μ

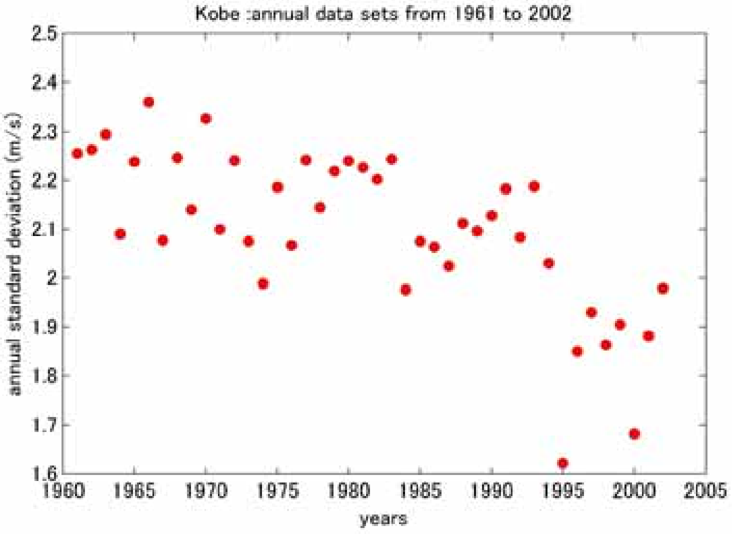
σ

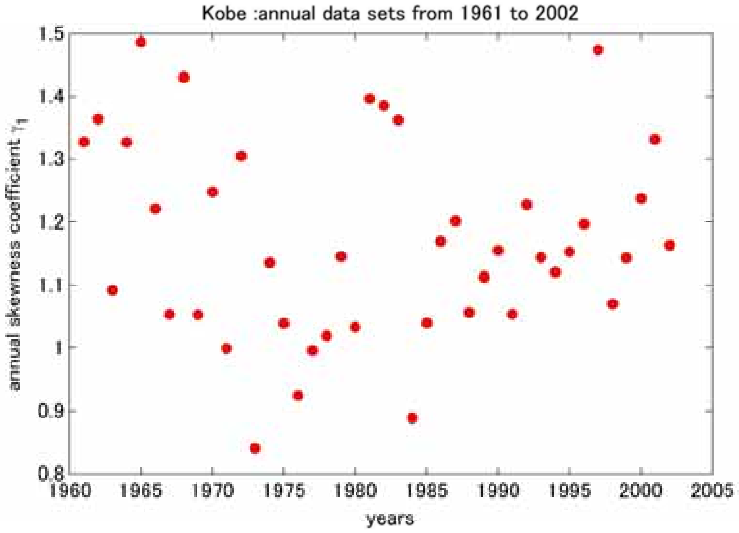
$\gamma_1 = \sqrt{\beta_1}$

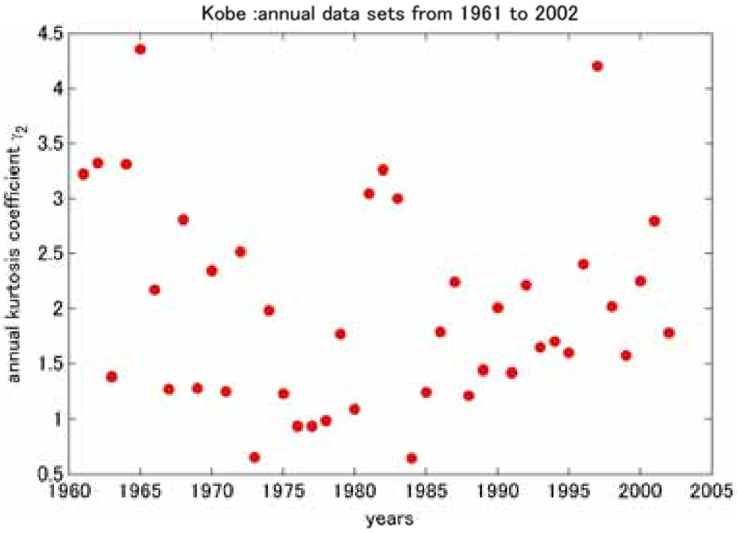
$\gamma_2 = \beta_2 - 3$

Kobe (1961~2002)

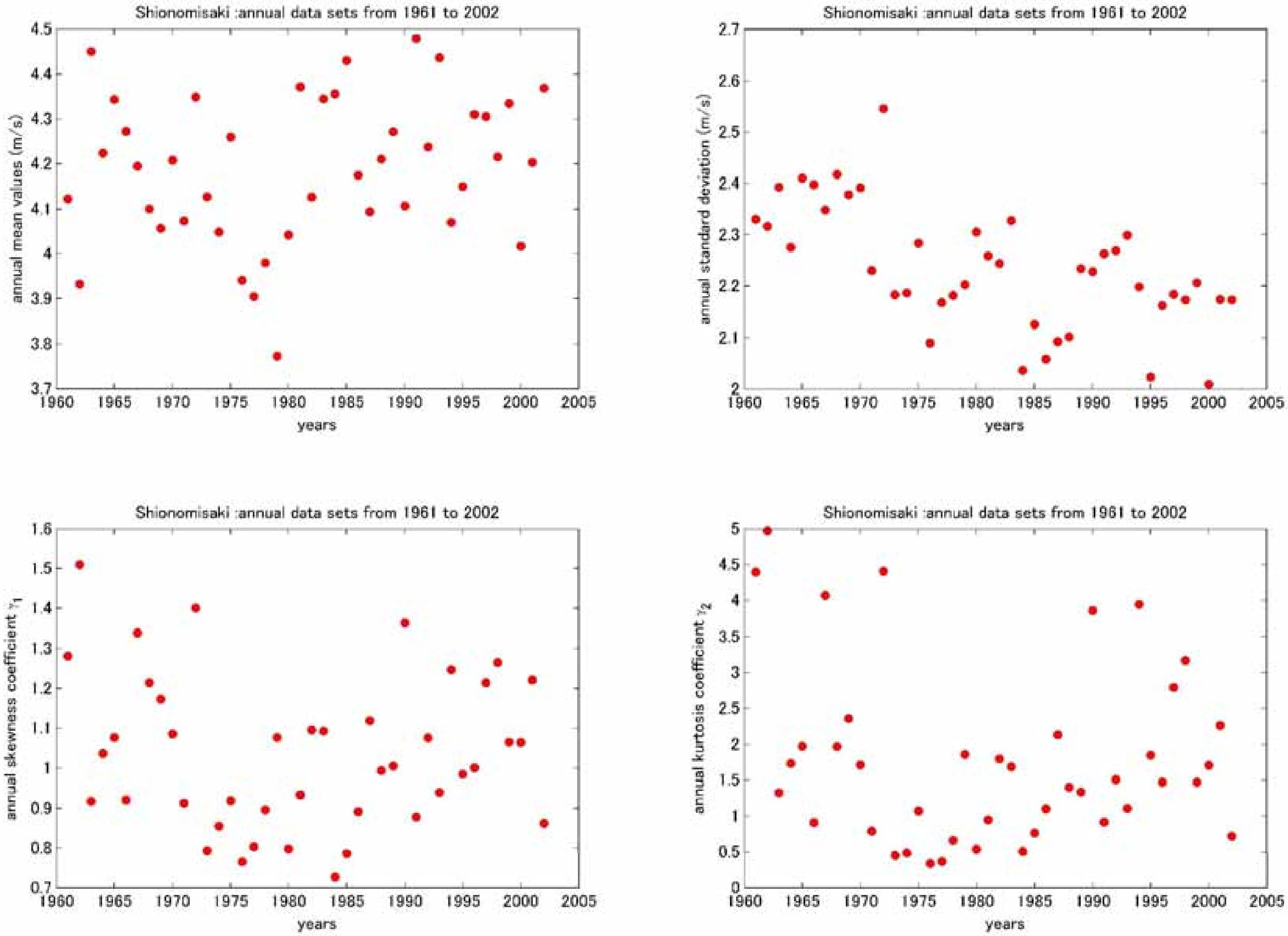

Kobe :annual data sets from 1961 to 2002


Kobe :annual data sets from 1961 to 2002

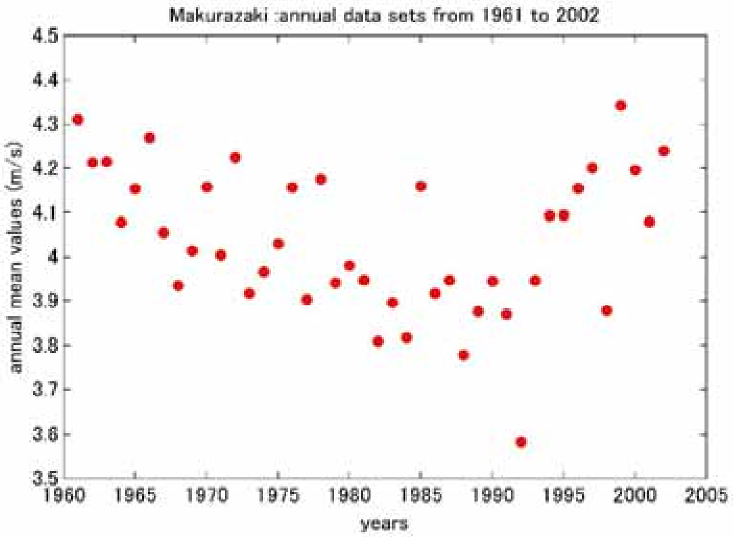

Kobe :annual data sets from 1961 to 2002

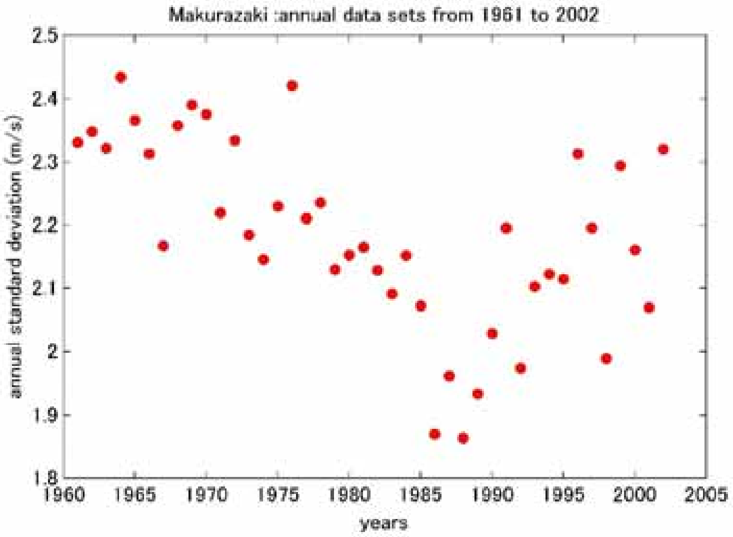

Kobe :annual data sets from 1961 to 2002

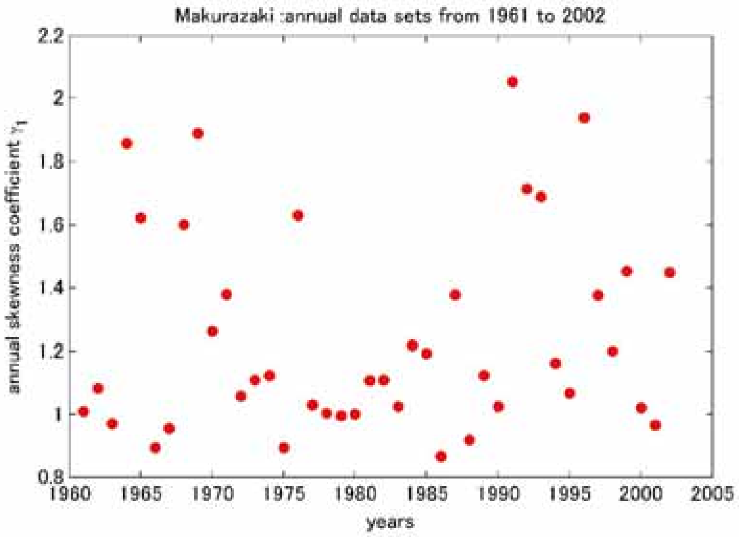
Shionomisaki (1961~2002)

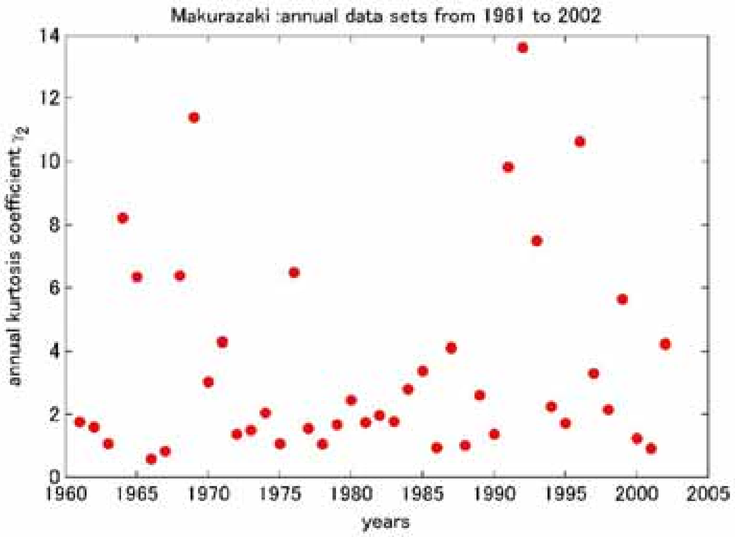


Makurazaki (1961~2002)

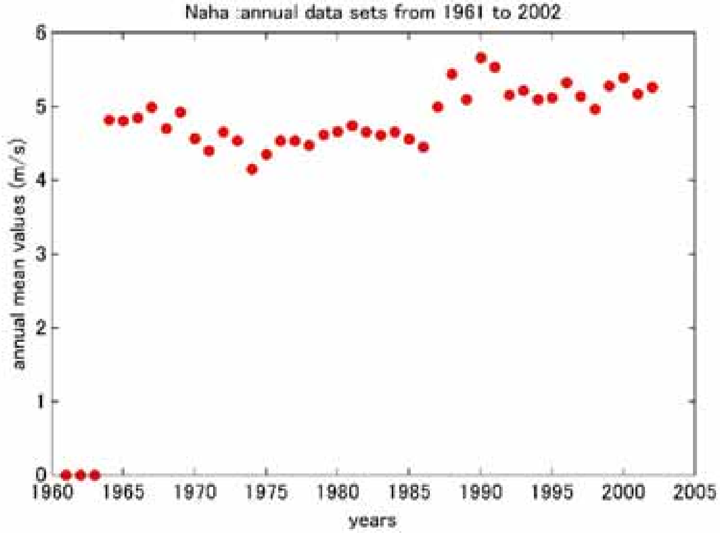


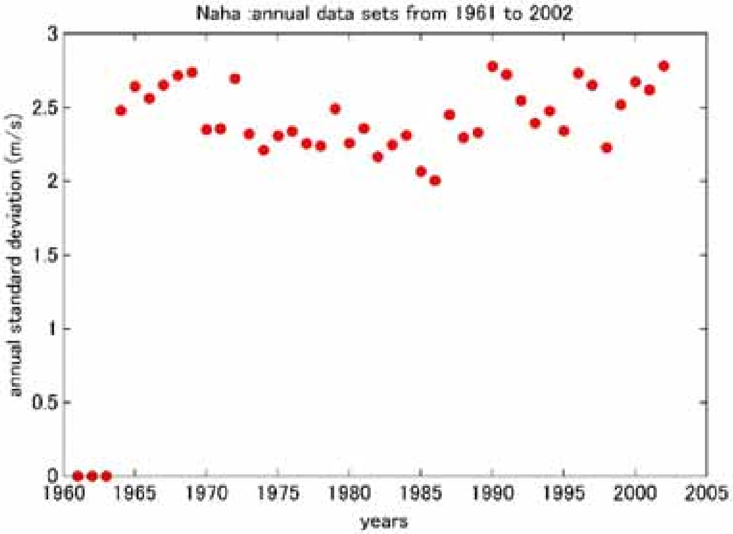


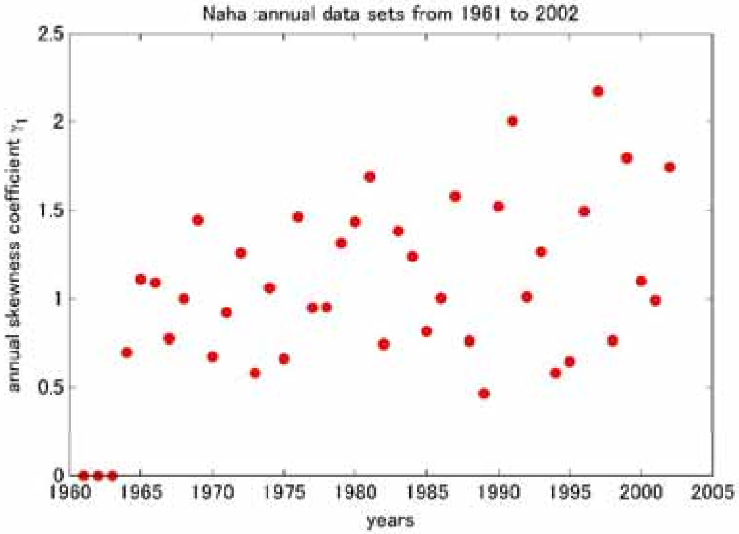


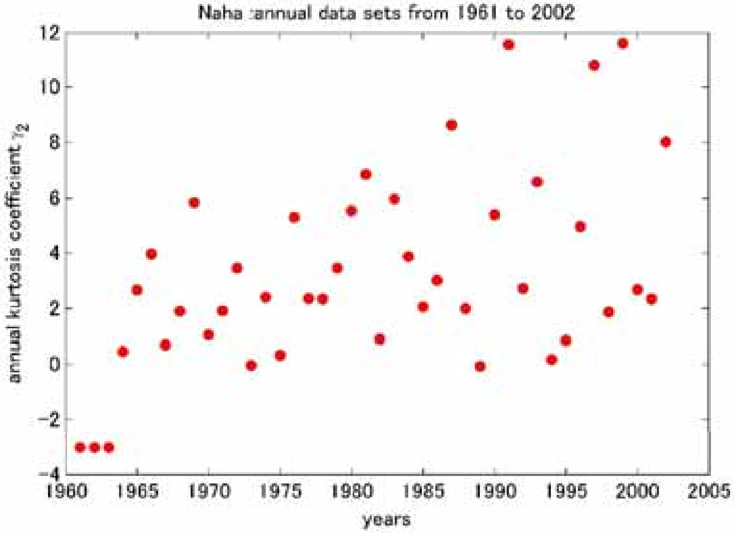


Naha (1964~2002)









09 practical application:
Annual maximum wind speed in Japan

4) Approximation of the annual wind speed distribution

Based on the probability integral transformation,

$$\Phi(z_\alpha) = F_{X_i}(x_\alpha) = F_{X_i}\left(g(z_\alpha)\right) = \alpha$$

$$x_\alpha = g(z_\alpha) = a + bz_\alpha + cz_\alpha^2 + dz_\alpha^3$$

$$\Rightarrow \quad F_{X_i}(x_\alpha) = \Phi\left(g^{-1}(x_\alpha)\right)$$

The coefficients *a*,*b*,*c* and *d* can be estimated from the given basic statistics of annual wind speed, i.e. mean, standard deviation, skewness and kurtosis (Choi & Kanda 2003).

H. Choi and J. Kanda (2003), Translation Method: a historical review and its application to simulation on non-Gaussian stationary processes, *Wind and Struct*., 6(5), p357~386

5) Estimation by Monte Carlo Simulation (MCS)

5.1) Simulation methods

① Based on the Spectral representation theorem for stationary stochastic process

→Using a given spectral density function, discrete stationary stochastic process is simulated.

→time consuming method

② Based on equivalent i.i.d. random sequence (EQRS)

→A stochastic process, which can be approximated by Poisson process, is modeled as an i.i.d. random sequence having same quantile function.

→time effective method

5.2) Required information for MCS based on EQRS

① m : the number of Independent rv

→ approximated by mean zero crossing rate (Normal process)

From Rice theorem and Poisson approximation, normal quantile function is given as follows:

$$z_\alpha = \sqrt{2\left\{\log \mu_0 T + y_T\right\}}$$

in whcih μ_0 : mean zero crossing rate, $y_T = -\log(-\log \alpha)$

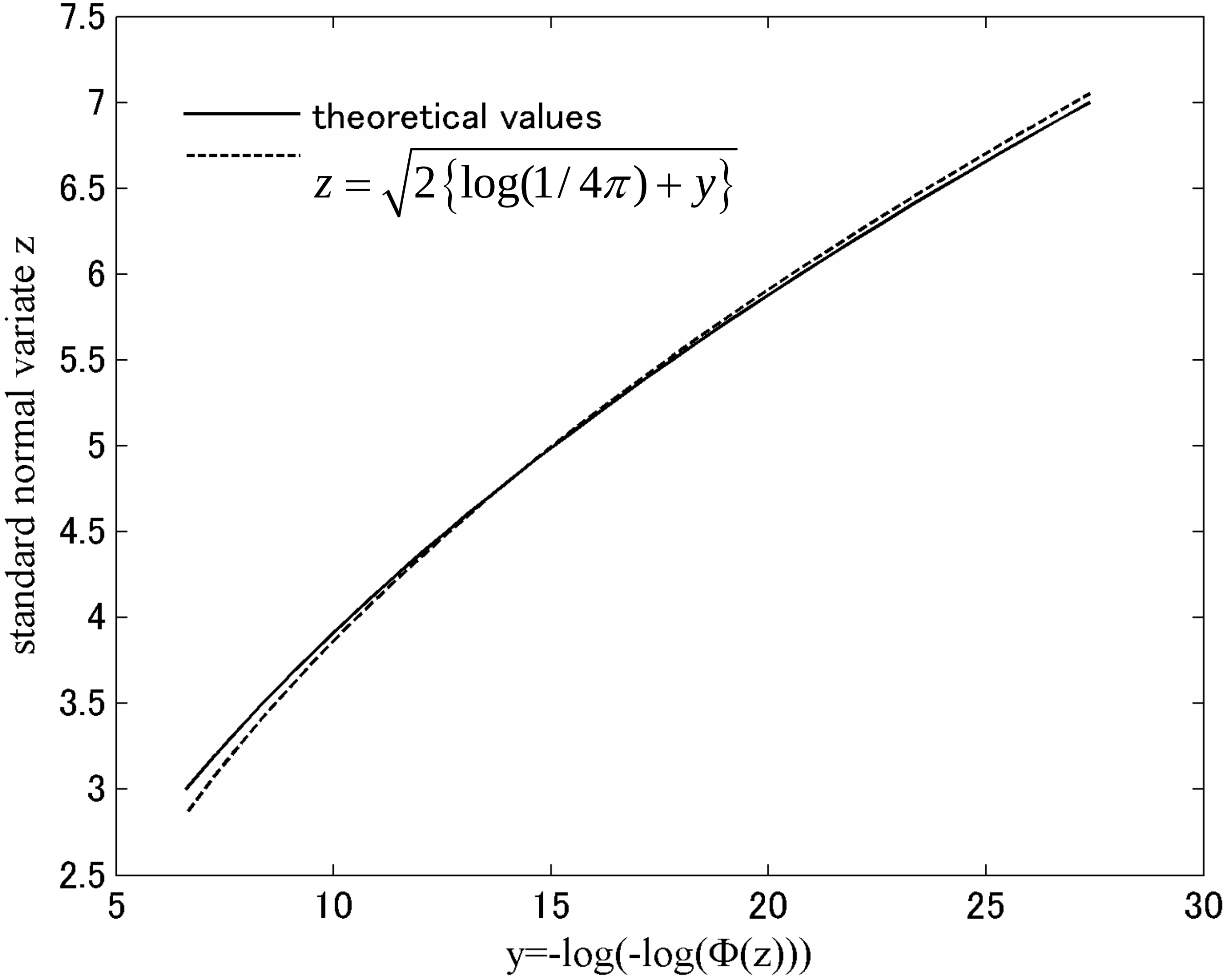
From $\Phi^m(z_\alpha)$,

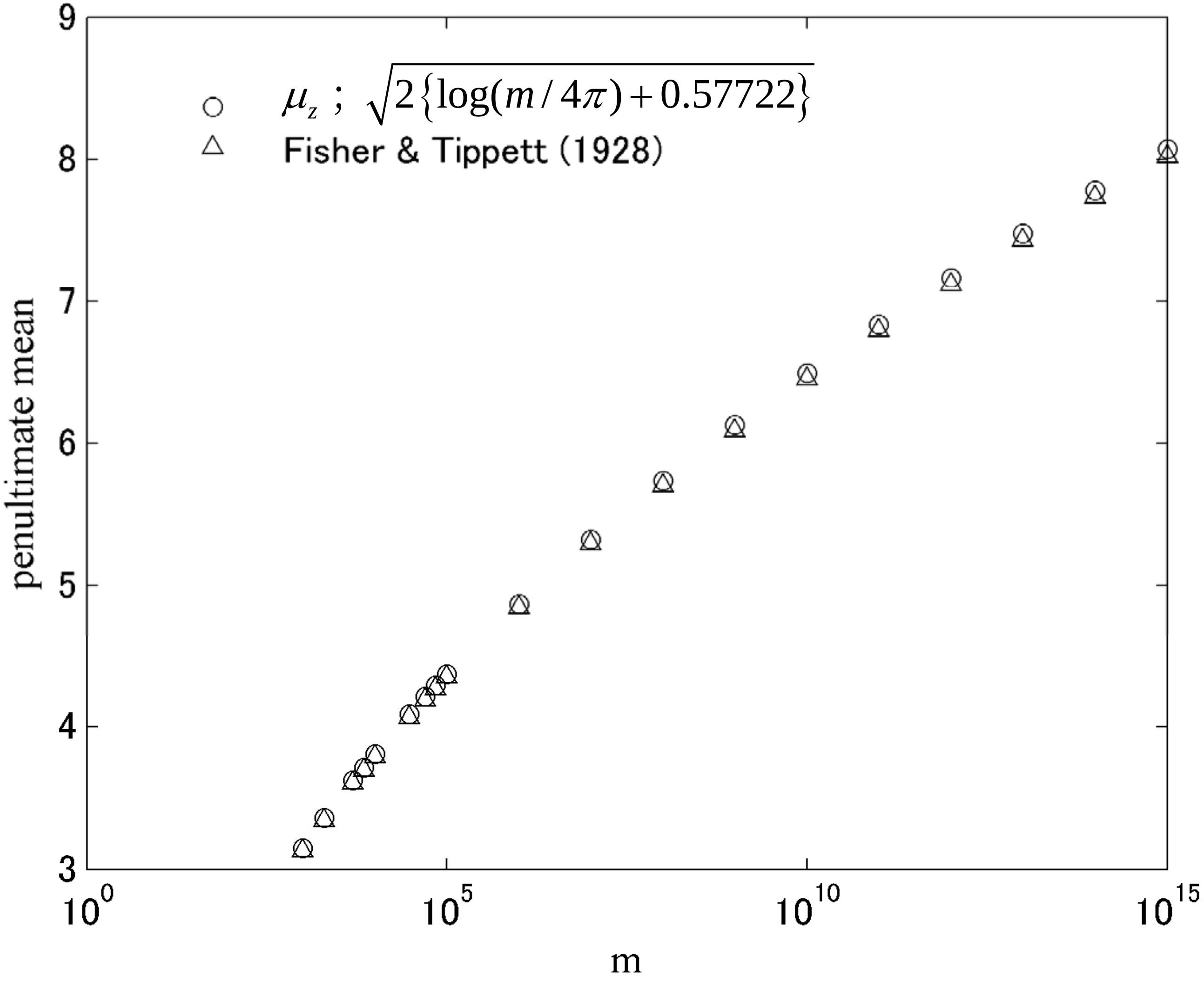
$$z_\alpha \; ; \; \sqrt{2\left\{\log\left(m/4\pi\right) + y_T\right\}} \quad *$$

By comparison,

$$m \; ; \; 4\pi\mu_0 T$$

* Choi & Kanda (2004), A new method of the extreme value distributions based on the translation method, *Summaries of Tech. Papers of Ann. Meet. of AIJ*, Vol. B1, p23~24 (in Japanese)


theoretical values
$z = \sqrt{2\{\log(1/4\pi) + y\}}$
standard normal variate z
y=-log(-log(Φ(z)))
7.5
7
6.5
6
5.5
5
4.5
4
3.5
3
2.5
5
10
15
20
25
30


9
8
7
6
5
4
3
penultimate mean
10^0
10^5
10^{10}
10^{15}
m
○ μ_z ; $\sqrt{2\{\log(m/4\pi)+0.57722\}}$
△ Fisher & Tippett (1928)

(non-Normal process)

To Rice formula for the expected number of crossings, i.e.

$$\nu(x)=\int_0^\infty \dot{x}\cdot f(\dot{x}|\,X=x)d\dot{x}=f(x)\cdot\int_0^\infty \dot{x}\cdot f(\dot{x})d\dot{x}$$

applying translation function *g*(*z*)

$$\nu(x)=\frac{1}{|g'(z)|}\phi\left(g^{-1}(x)\right)\cdot\int_0^\infty\frac{\dot{x}\cdot\exp\left\{-\left(\dot{x}/\sigma_{\dot{x}}\cdot g'(z)\right)^2/2\right\}}{\sqrt{2\pi}\cdot\sigma_{\dot{x}}\cdot g'(z)}d\dot{x}$$

$$=\frac{\sigma_{\dot{x}}}{2\pi}\cdot\exp\left(-\frac{\{g^{-1}(x)\}^2}{2}\right)=\nu_0\cdot\exp\left(-\frac{\{g^{-1}(x)\}^2}{2}\right)\quad *$$

From Poisson approximation

$$x_\alpha=g(z_\alpha)=g\left(\sqrt{2\{\log(\nu_0T)+y_T\}}\right)$$

With the same manner

$$m\;;\;4\pi\nu_0T$$

* The distribution of *dx/dt* is assumed as normal distribution and the assumption is reasonable. *e.g.* H. Choi (1988), *Characteristics of natural wind for wind load estimation*, Master Thesis, Univ. of Tokyo (in Japanese)

Practical example (*T*=1 year)

- $\nu_0 T$ estimated from long term observation records in Tokyo (1985~1987, Choi 1988)

Height (m)	45	45	46	48	58
$\nu_0 T$	2883	2665	2545	2739	2307
Height (m)	63	79	93	187	
$\nu_0 T$	3124	2586	3115	2384	

$$m = 4\pi\nu_0 T = 4\pi\cdot(2300 \sim 3000) \rightarrow 30,000$$

5.2) Required information for MCS based on EQRS

② parent distribution function for each year

→Generalized bootstrap method＋

Translation method (Probability Integral Transform)

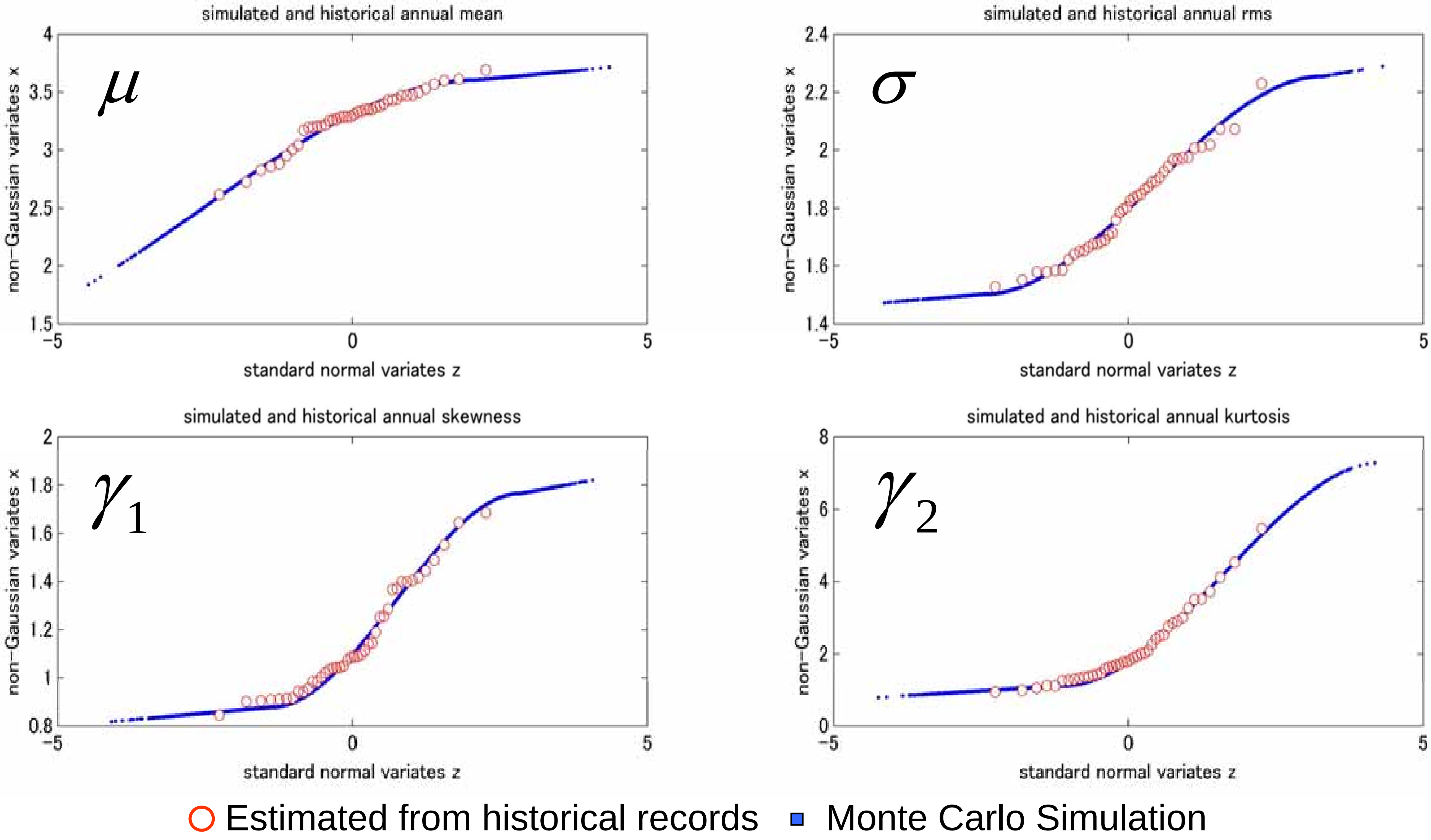
For the year *i*,

$$F_i(x = g_i(z)) = \Phi(z),\ g_i(z) = a_i + b_i z + c_i z^2 + d_i z^3$$

$$\{a,b,c,d\}_i = \Psi(\mu_i, \sigma_i, \gamma_{1i}, \gamma_{2i})$$

$$\max(x_1, \ldots, x_m)_{i=1,\ldots,n} = g_i(\max(z_1, \ldots, z_m)_{i=1,\ldots,n})$$

Example : Tokyo (1961~2002)



Correlation between the basic statistics (Tokyo)

	μ	σ	γ_1	γ_2
μ	1.000	0.242	-0.372	-0.315
σ	0.242	1.000	0.541	0.433
γ_1	-0.372	0.541	1.000	0.946
γ_2	-0.315	0.433	0.946	1.000

Such correlation characteristics between the basic statistics are regenerated by Cholesky decomposition of correlation matrix.

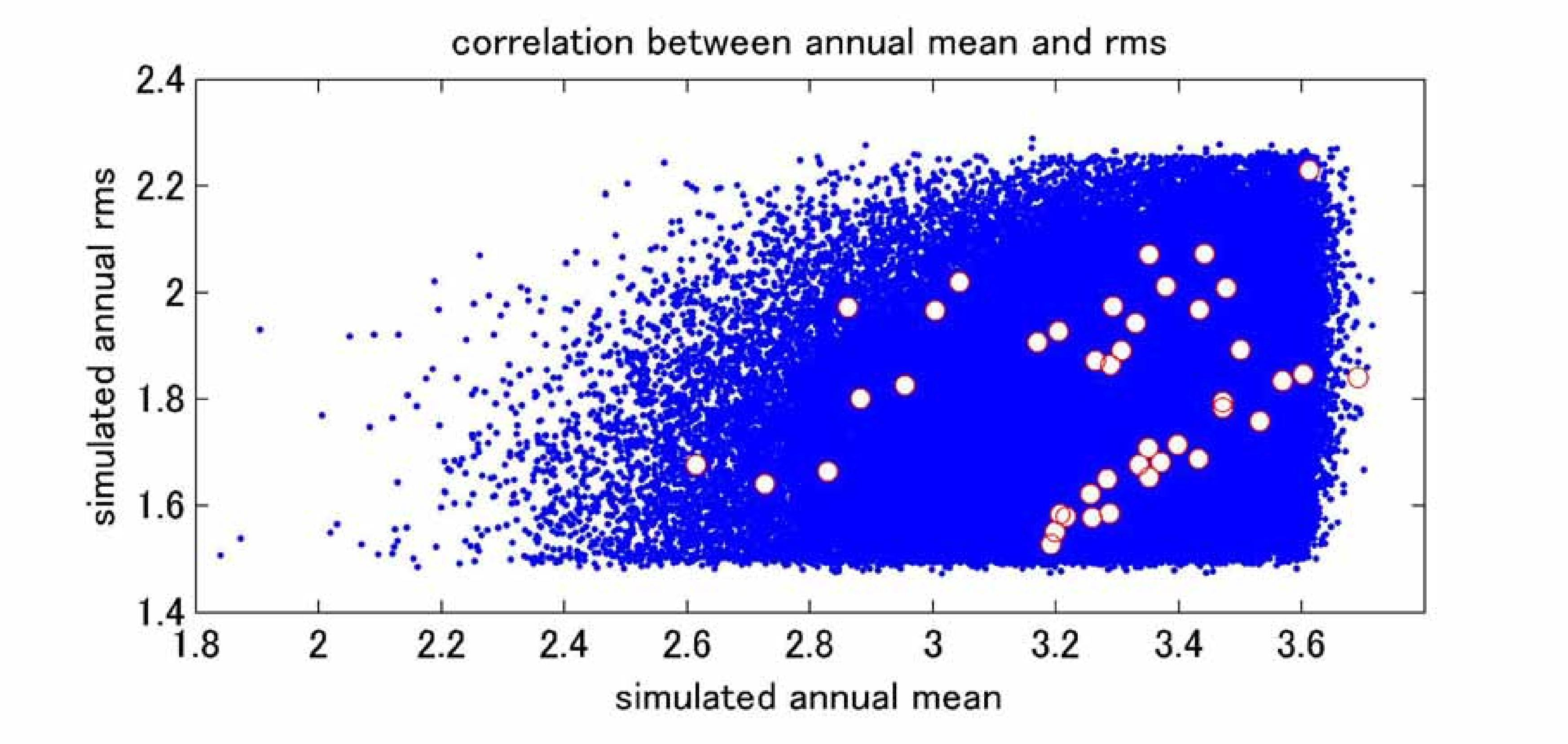
Regeneration of correlation characteristics

correlation coefficients of simulated ones and given values

	μ	σ	γ_1	γ_2
μ	1.000	0.242 (0.242)	-0.389 (-0.372)	-0.343 (-0.315)
σ	0.242 (0.242)	1.000	0.565 (0.541)	0.476 (0.433)
γ_1	-0.389 (-0.372)	0.565 (0.541)	1.000	0.958 (0.946)
γ_2	-0.343 (-0.315)	0.476 (0.433)	0.958 (0.946)	1.000

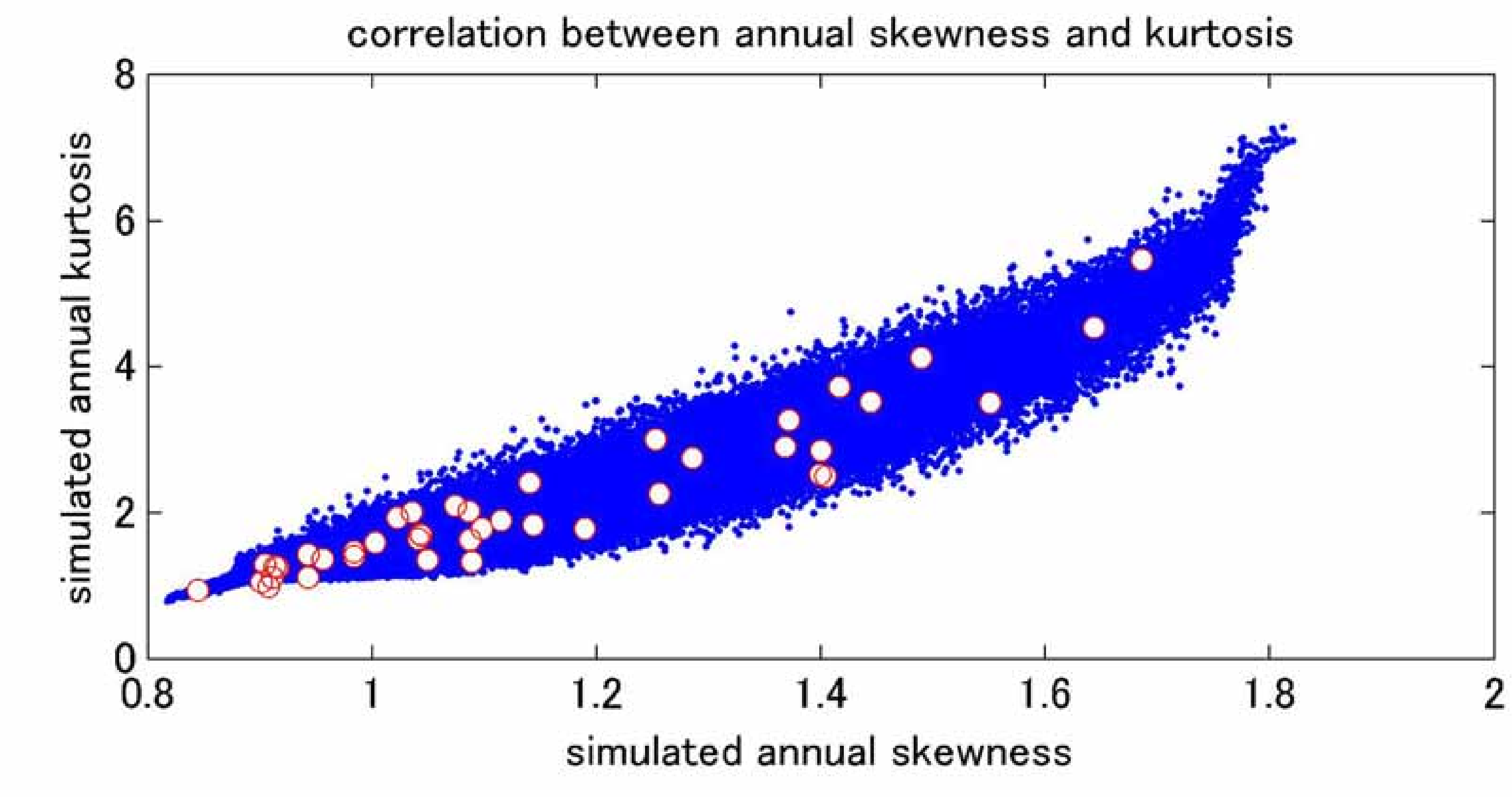
(･): given correlation coefficient

No. of Simulation：n=100 year x 1000 times

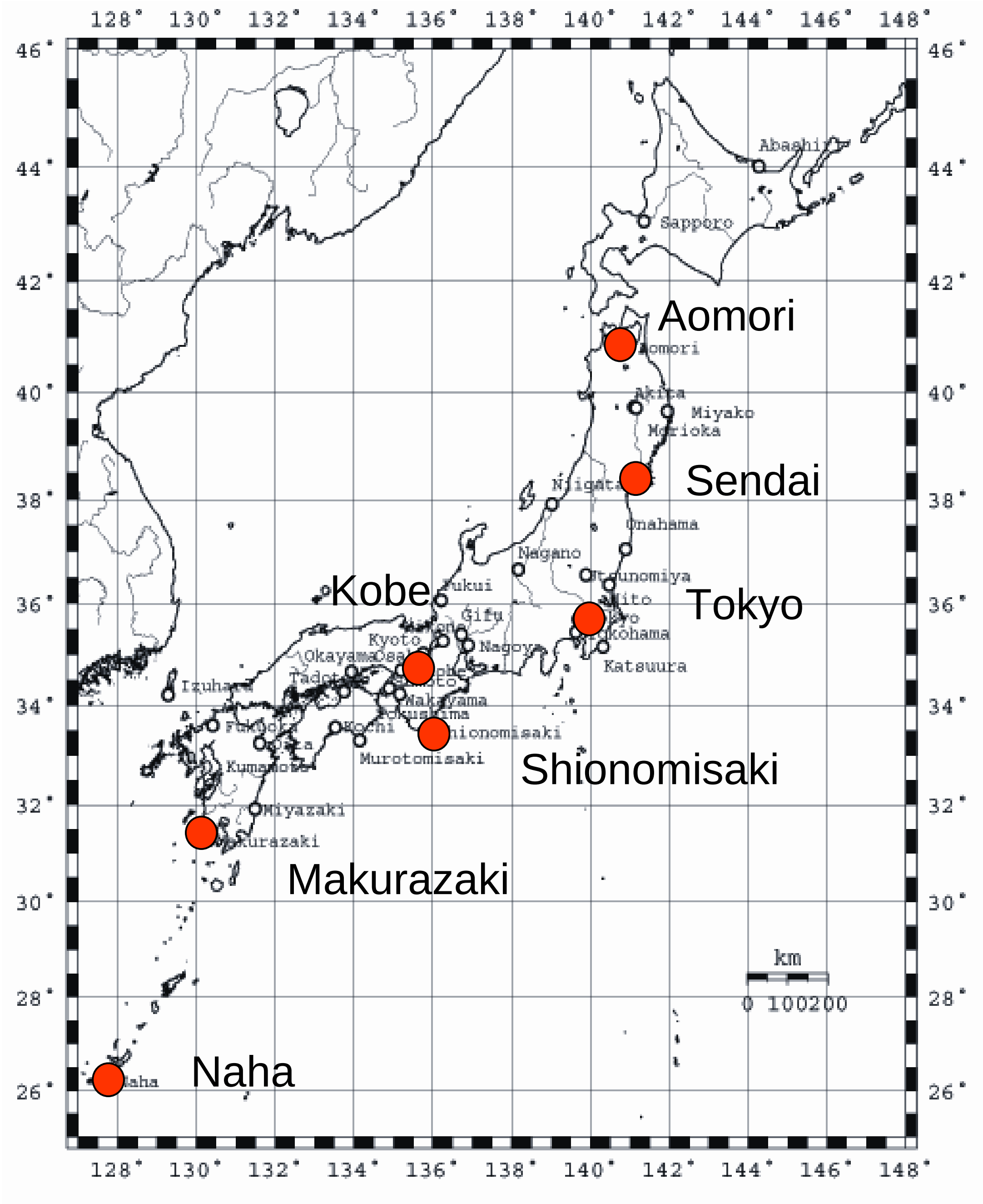


○：annual mean and standard dev. from historical records

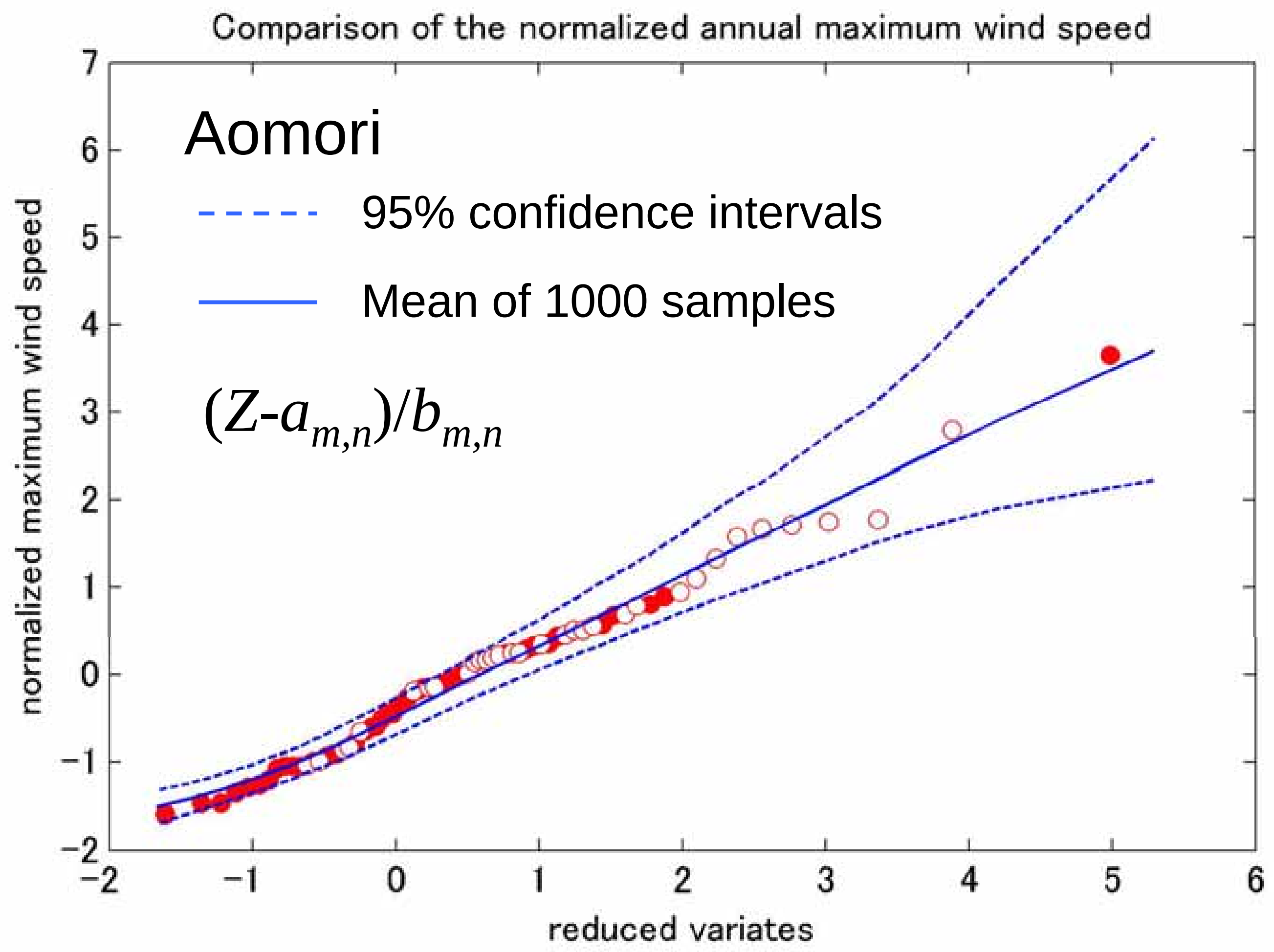
No. of Simulation：n=100 years x 1000 times



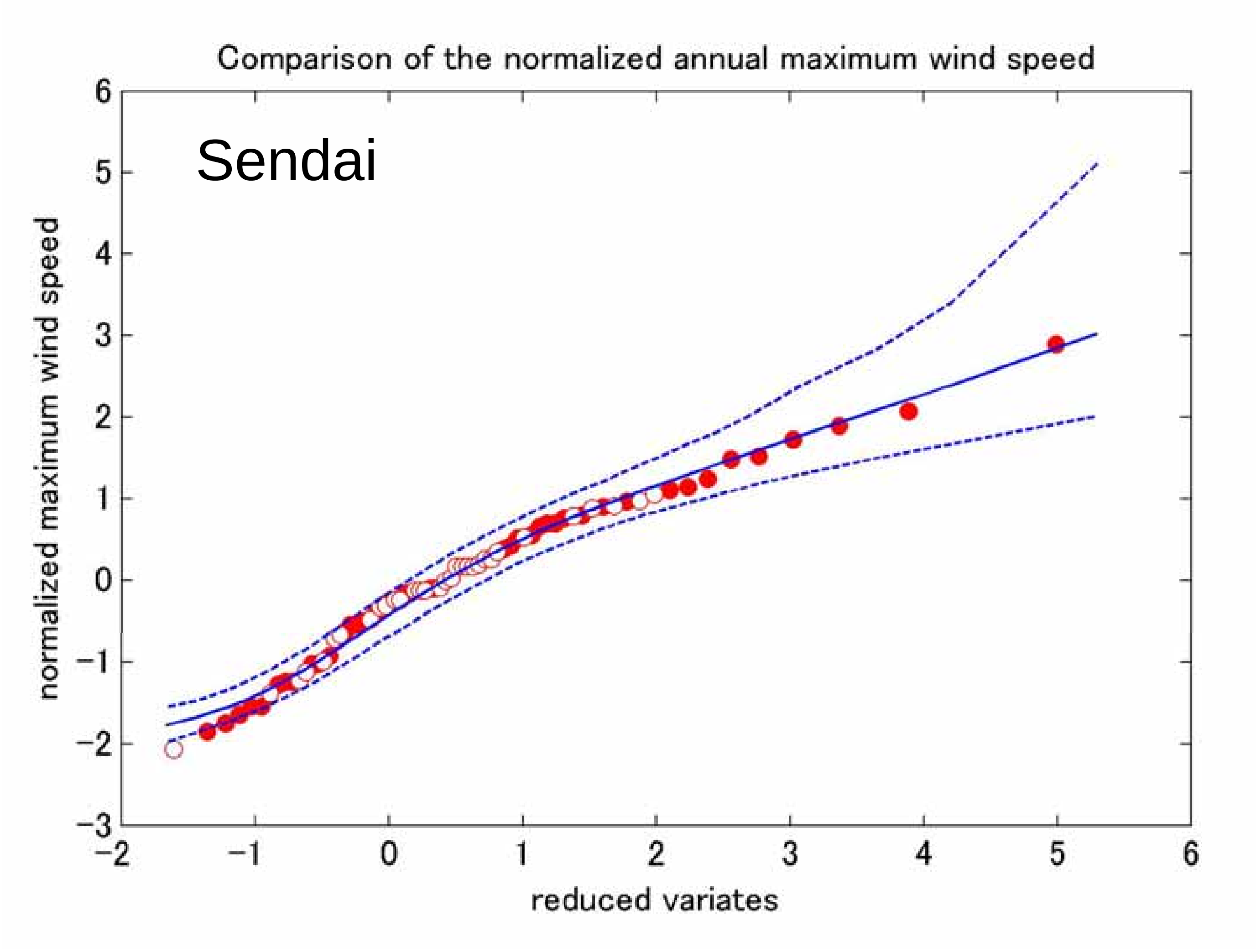
○：skewness and kurtosis from historical records

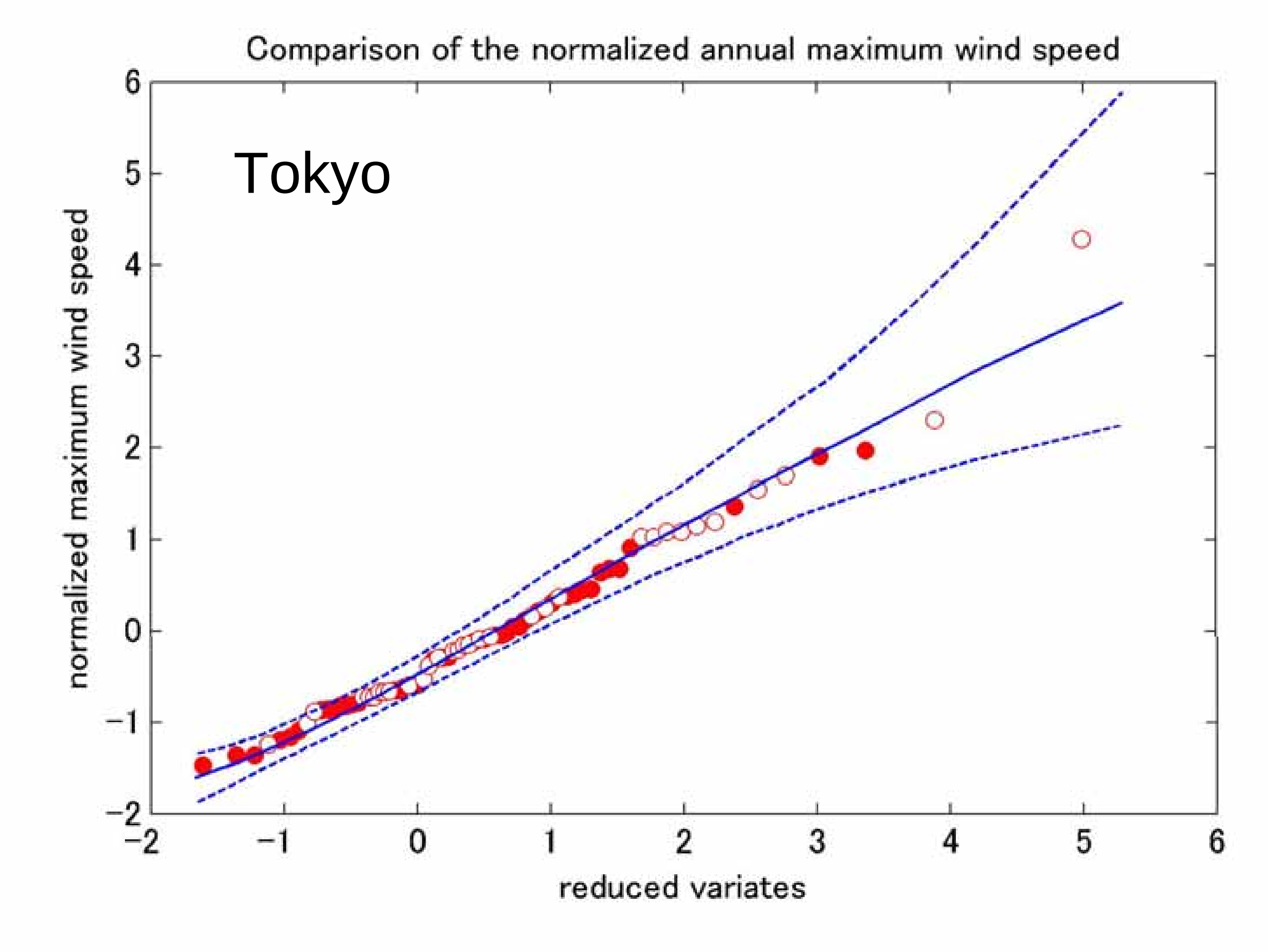

Aomori
Sendai
Tokyo
Kobe
Shionomisaki
Makurazaki
Naha
Abashiri
Sapporo
Akita
Miyako
Morioka
Niigata
Onahama
Nagano
Utsunomiya
Mito
Fukui
Gifu
Kyoto
Nagoya
Yokohama
Okayama
Osaka
Tadotsu
Izuhara
Wakayama
Tokushima
Katsuura
Fukuoka
Oita
Kochi
Murotomisaki
Kumamoto
Miyazaki
km
0 100200
128°
130°
132°
134°
136°
138°
140°
142°
144°
146°
148°
46°
44°
42°
40°
38°
36°
34°
32°
30°
28°
26°

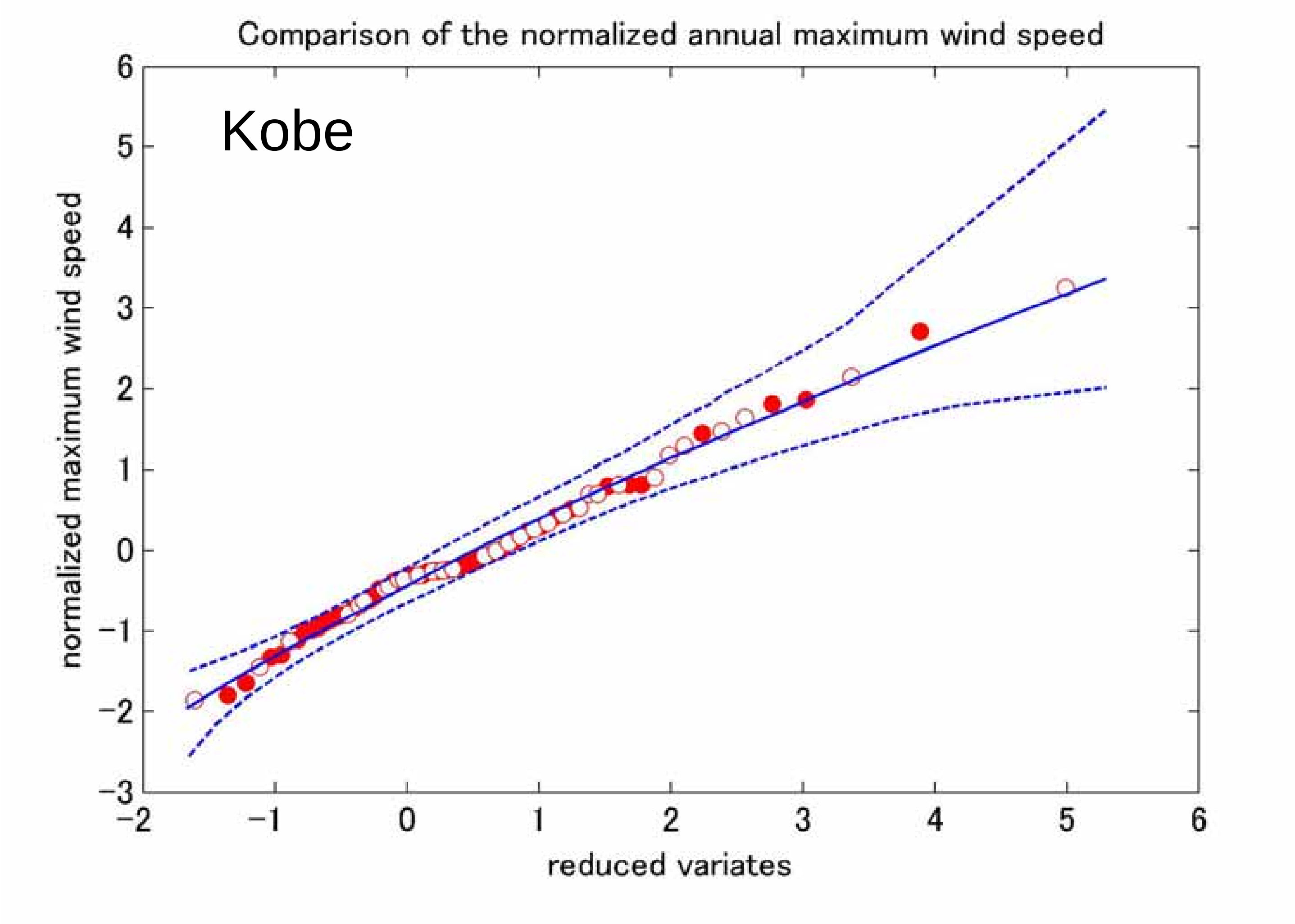
6) Comparison of the quantile functions from MCS and historical records in normalized form

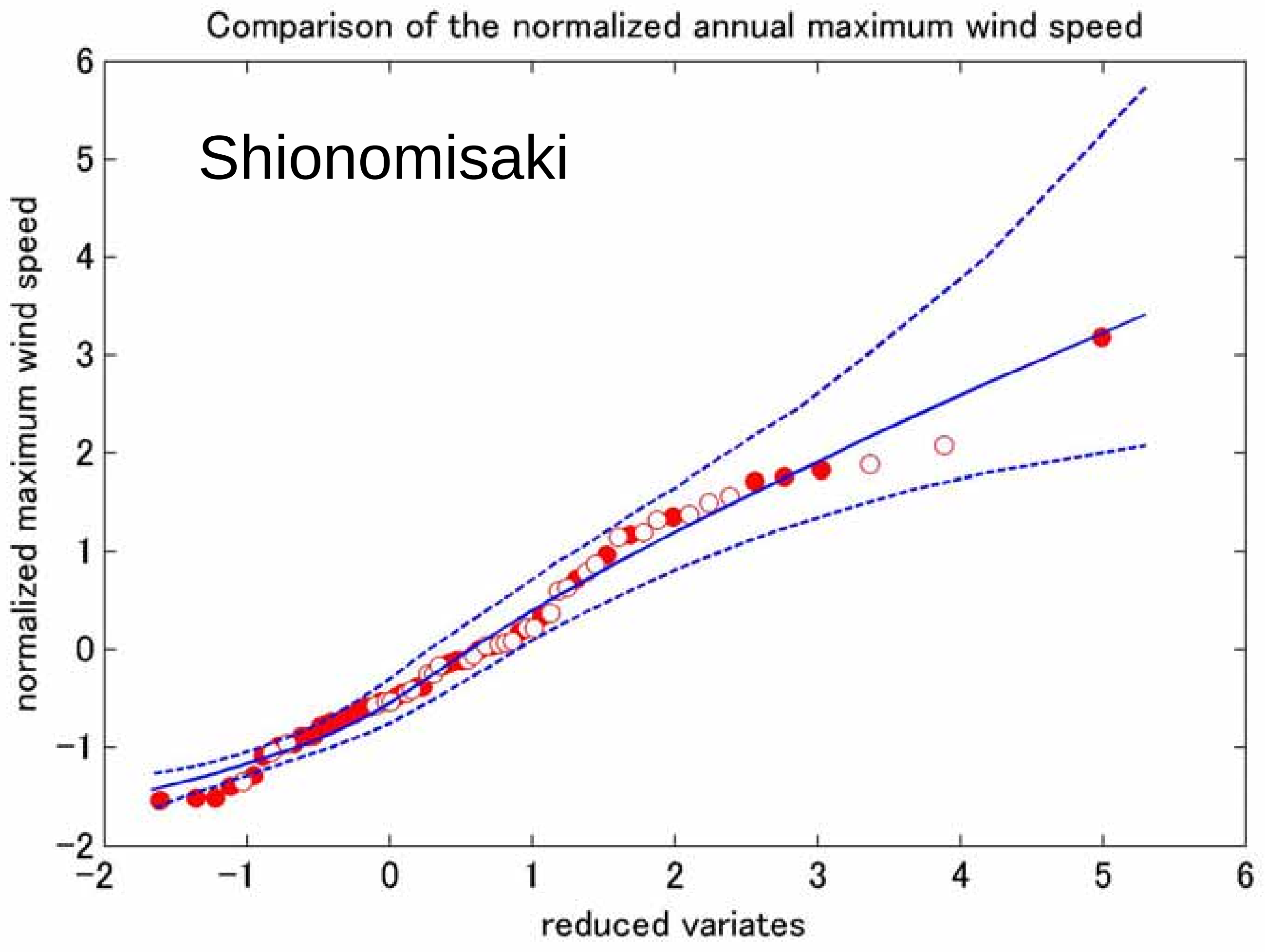


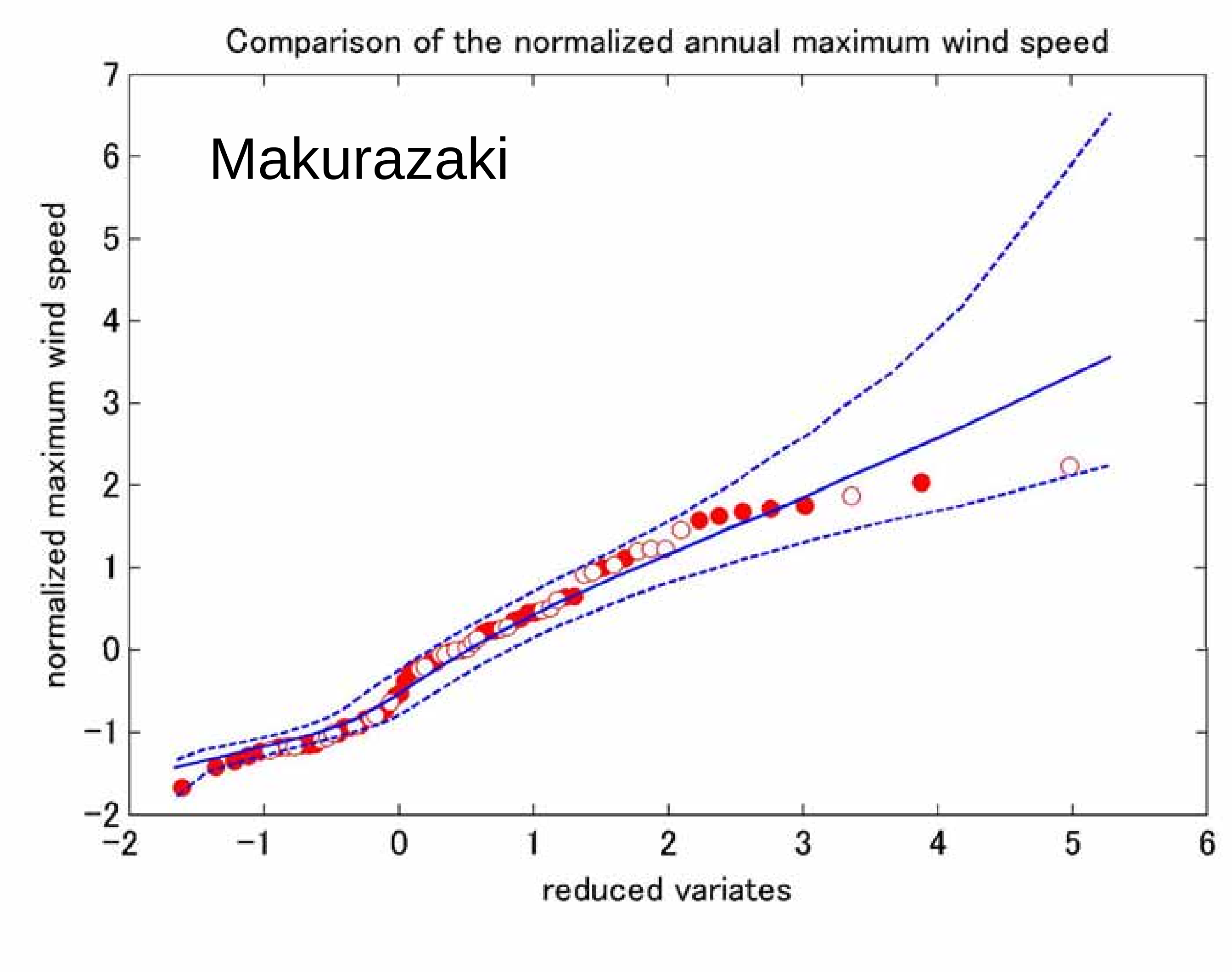
○ : before 1961 ● : after 1962

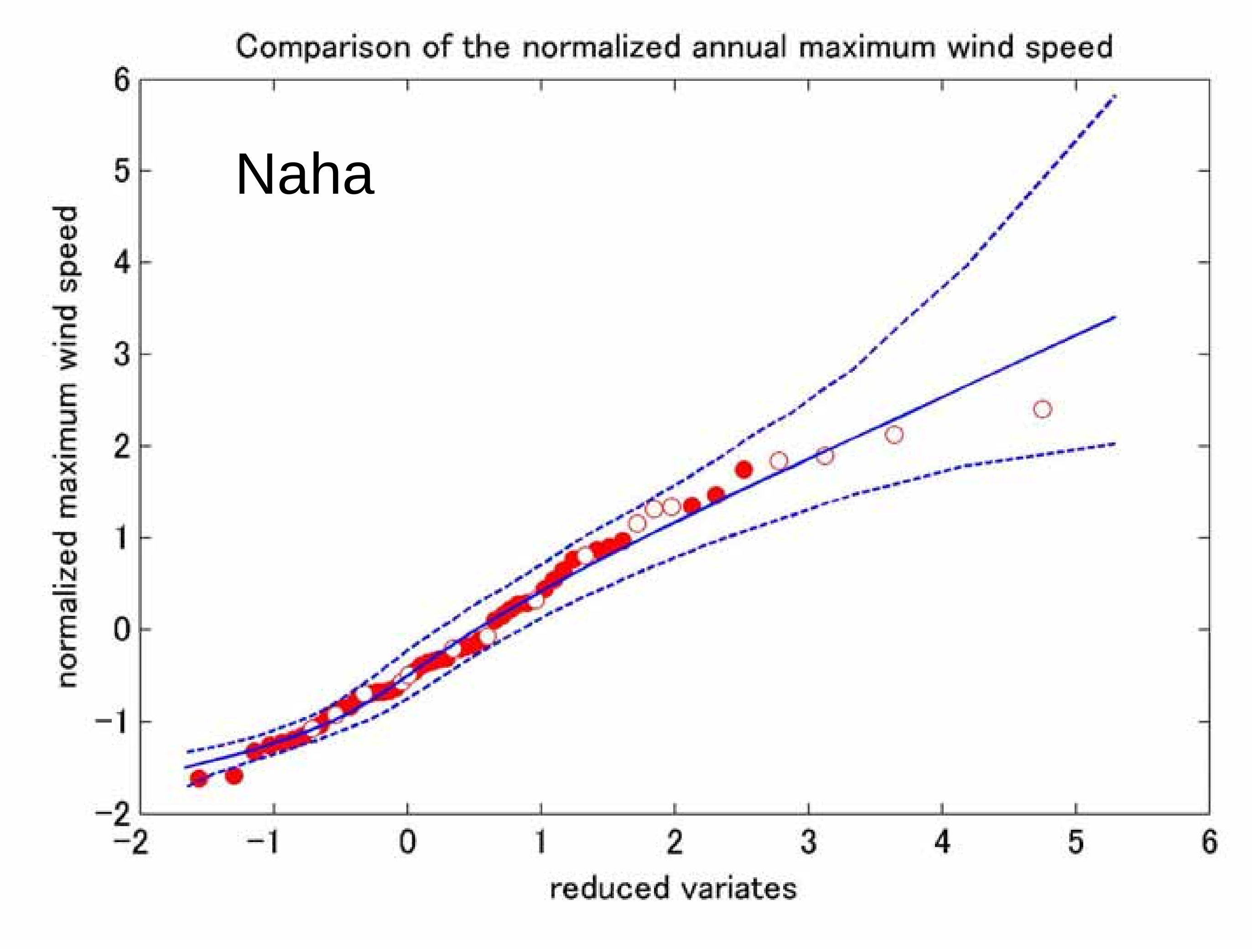
Comparison of the normalized annual maximum wind speed
Sendai
normalized maximum wind speed
reduced variates
6
5
4
3
2
1
0
−1
−2
−3
−2
−1
0
1
2
3
4
5
6


Comparison of the normalized annual maximum wind speed
Tokyo
normalized maximum wind speed
6
5
4
3
2
1
0
−1
−2
−2
−1
0
1
2
3
4
5
6
reduced variates

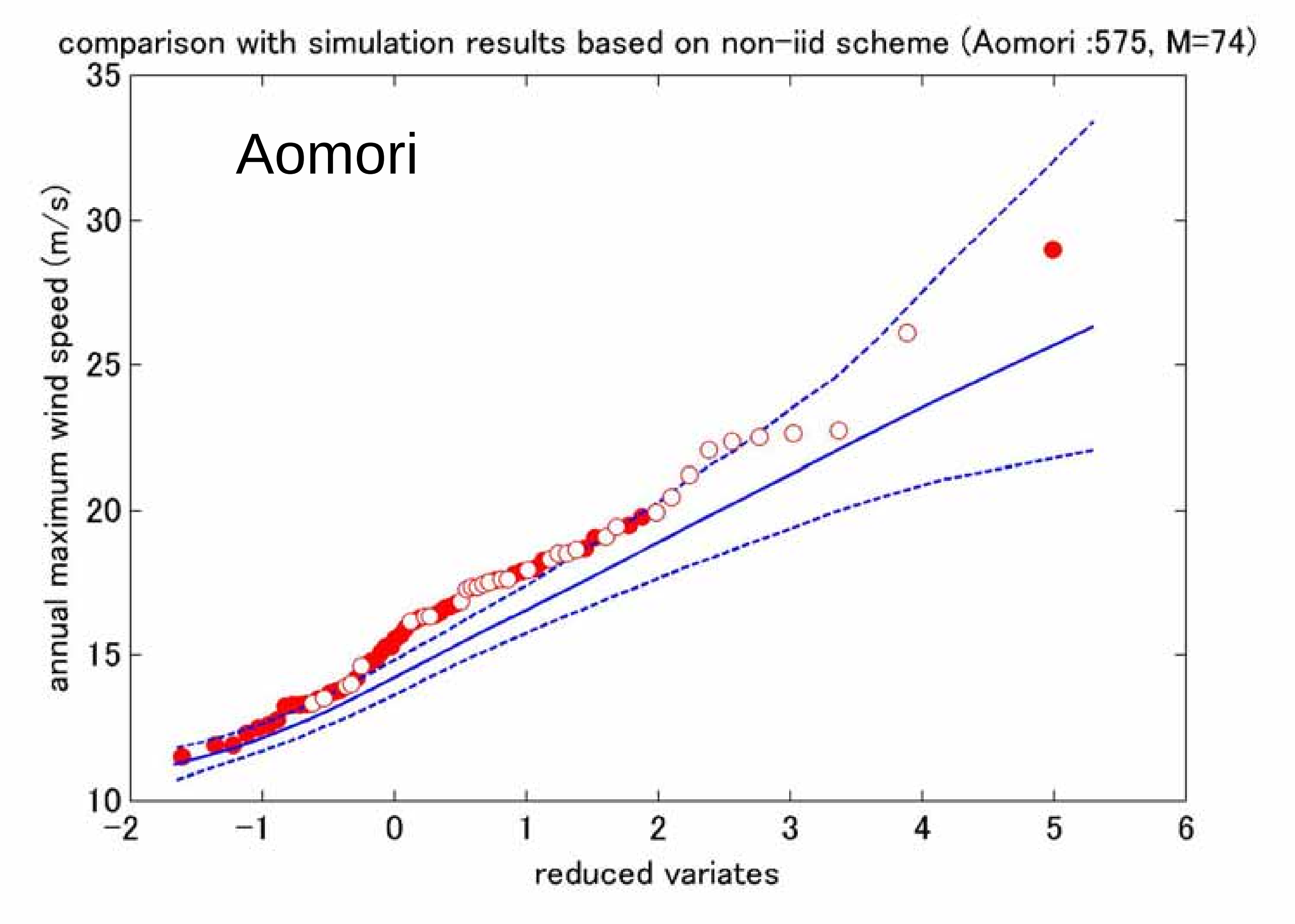

Comparison of the normalized annual maximum wind speed
Kobe
normalized maximum wind speed
reduced variates
6
5
4
3
2
1
0
−1
−2
−3
−2
−1
0
1
2
3
4
5
6

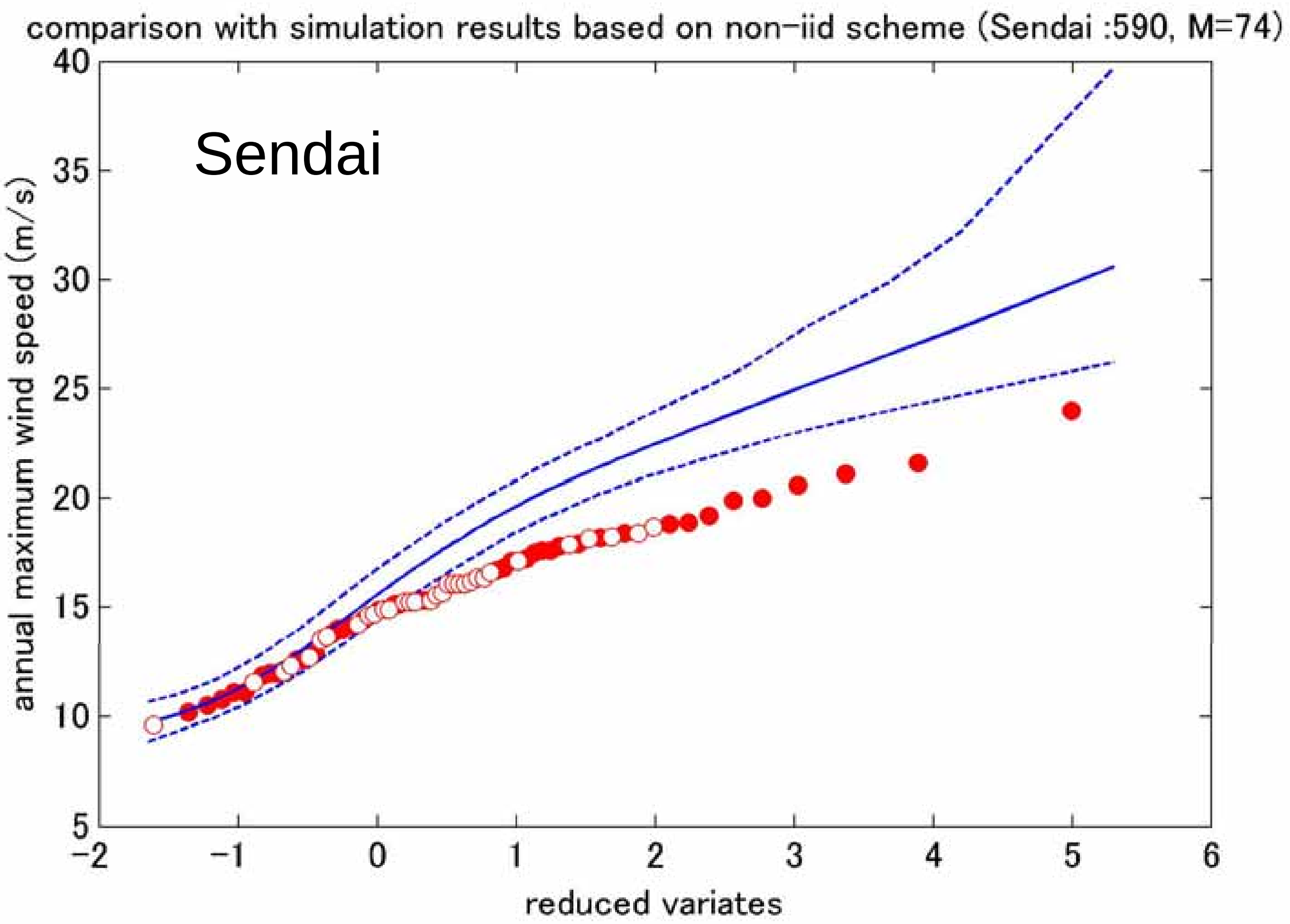

Comparison of the normalized annual maximum wind speed
Shionomisaki
normalized maximum wind speed
reduced variates
6
5
4
3
2
1
0
−1
−2
−2
−1
0
1
2
3
4
5
6


Comparison of the normalized annual maximum wind speed
Makurazaki
normalized maximum wind speed
reduced variates
7
6
5
4
3
2
1
0
−1
−2
−2
−1
0
1
2
3
4
5
6

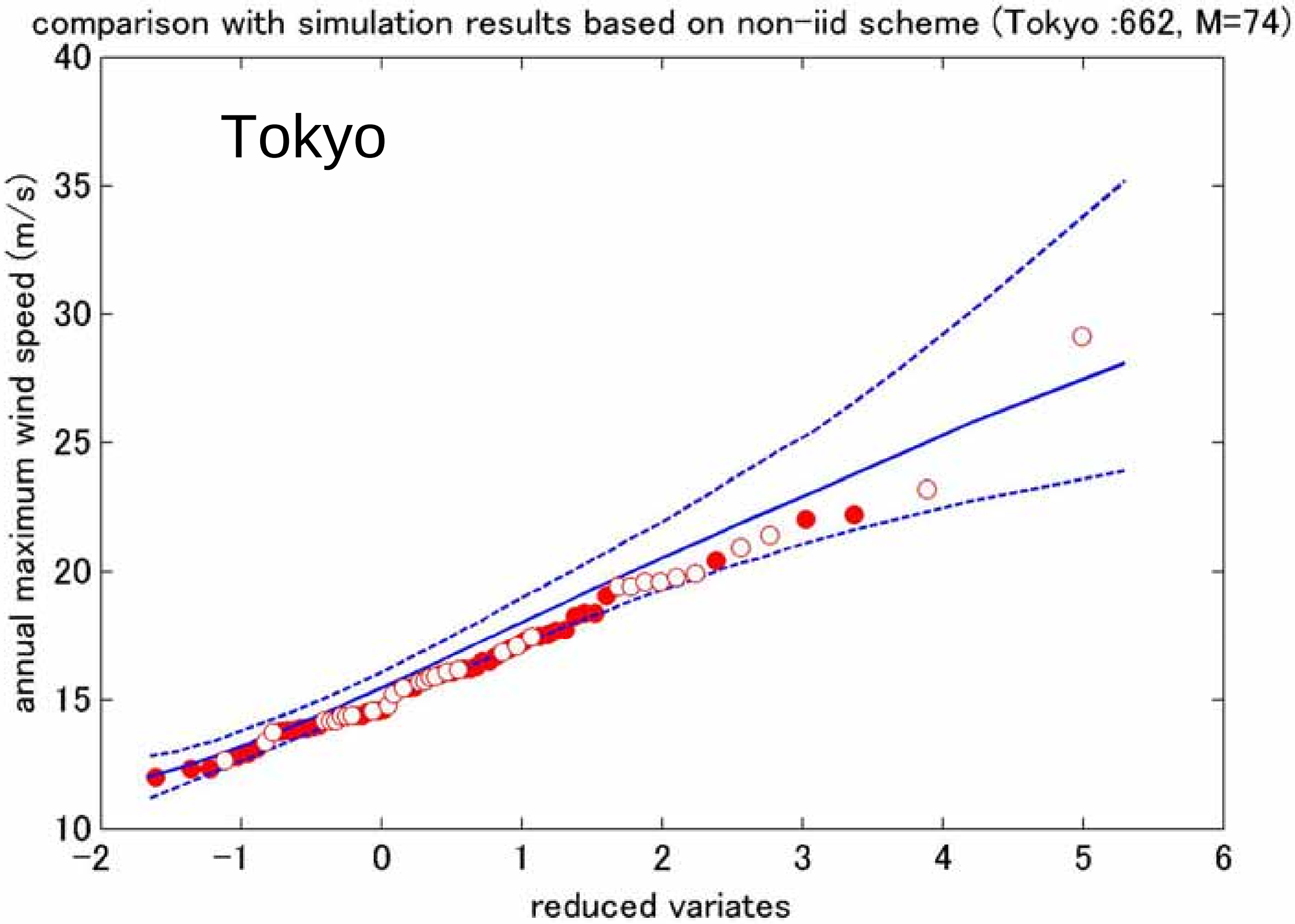

Comparison of the normalized annual maximum wind speed
Naha
normalized maximum wind speed
reduced variates
6
5
4
3
2
1
0
−1
−2
−2
−1
0
1
2
3
4
5
6

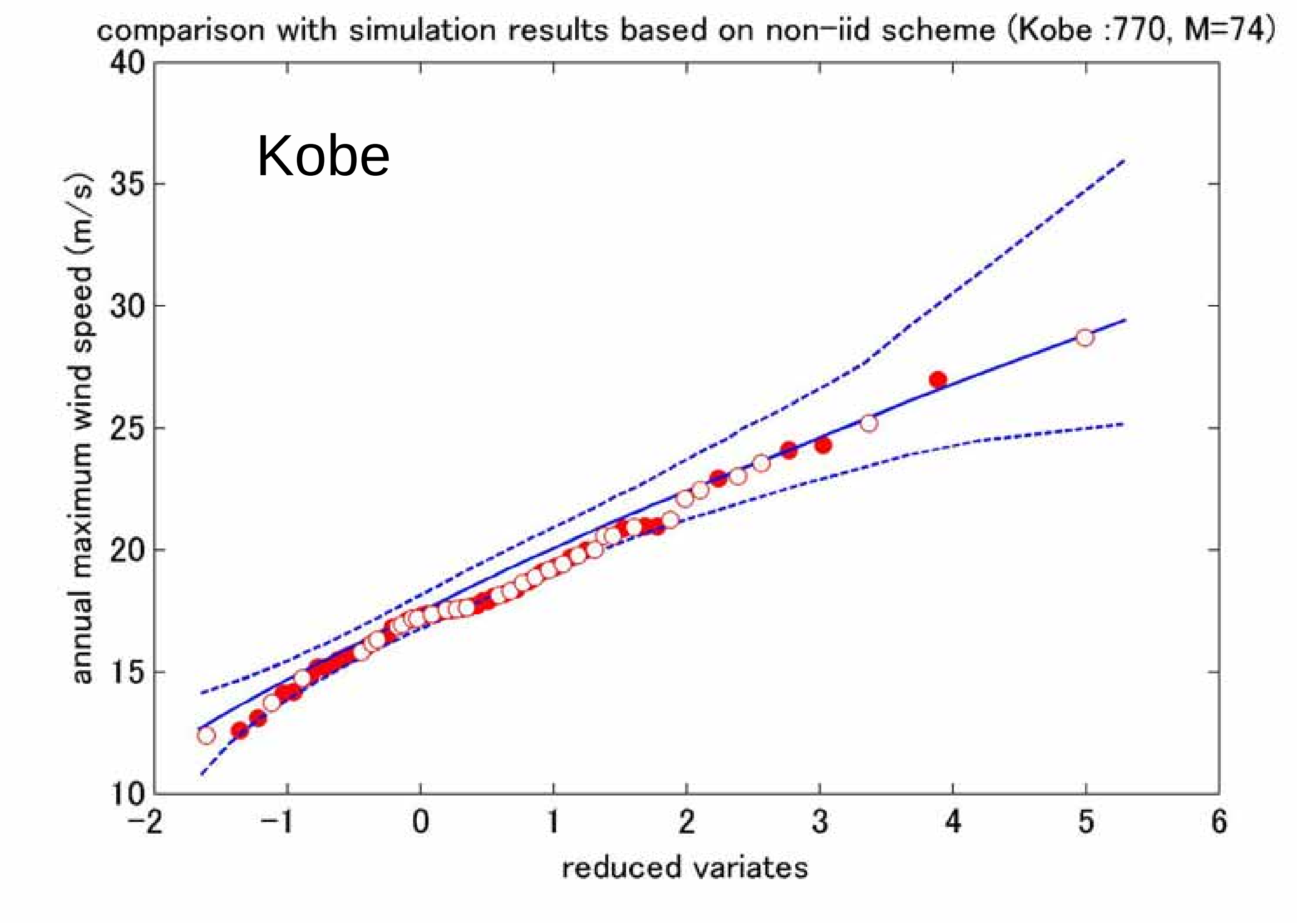
7) Comparison of the quantile functions from MCS and historical records in full scale

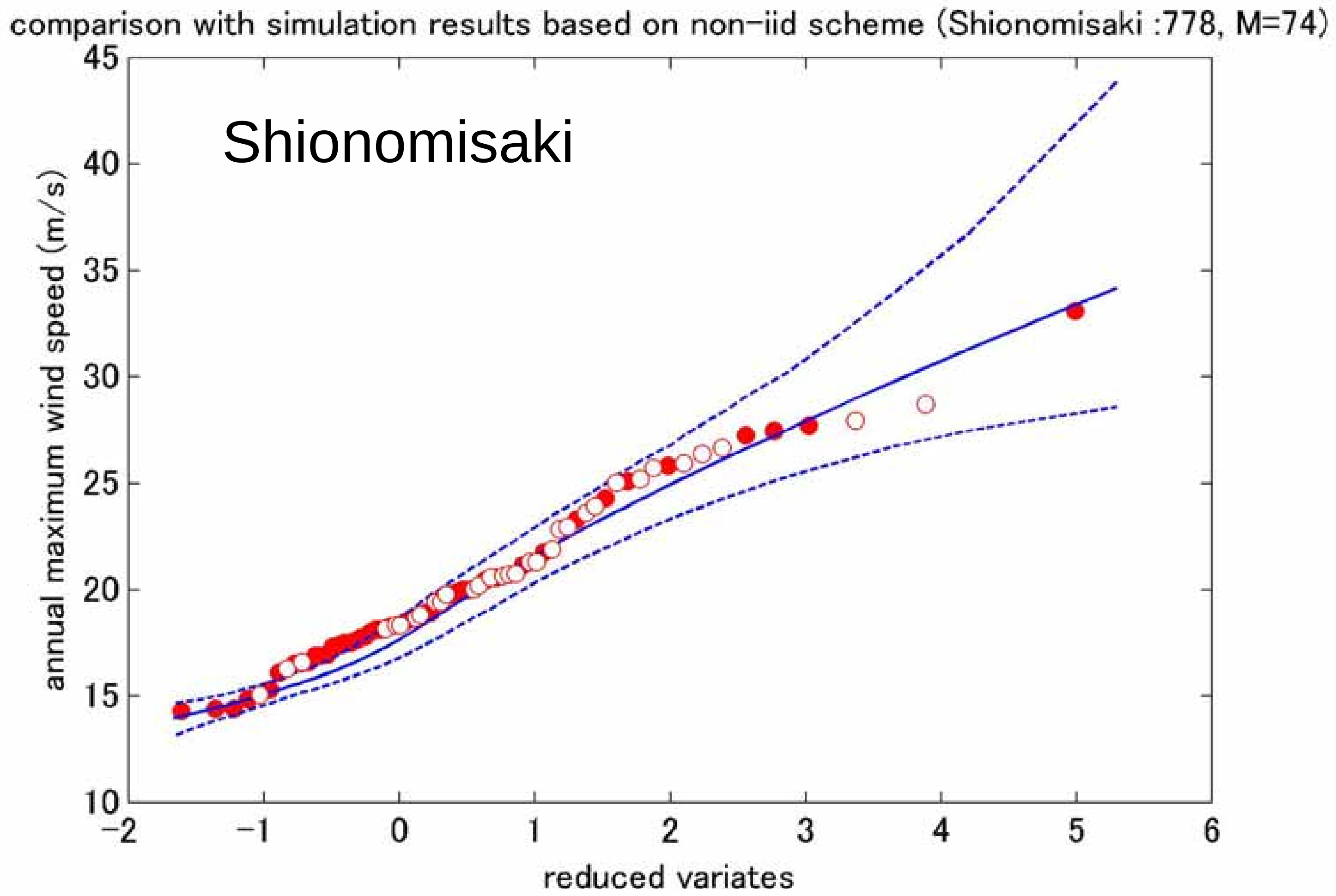



comparison with simulation results based on non-iid scheme (Sendai :590, M=74)
annual maximum wind speed (m/s)
40
35
30
25
20
15
10
5
−2
−1
0
1
2
3
4
5
6
reduced variates

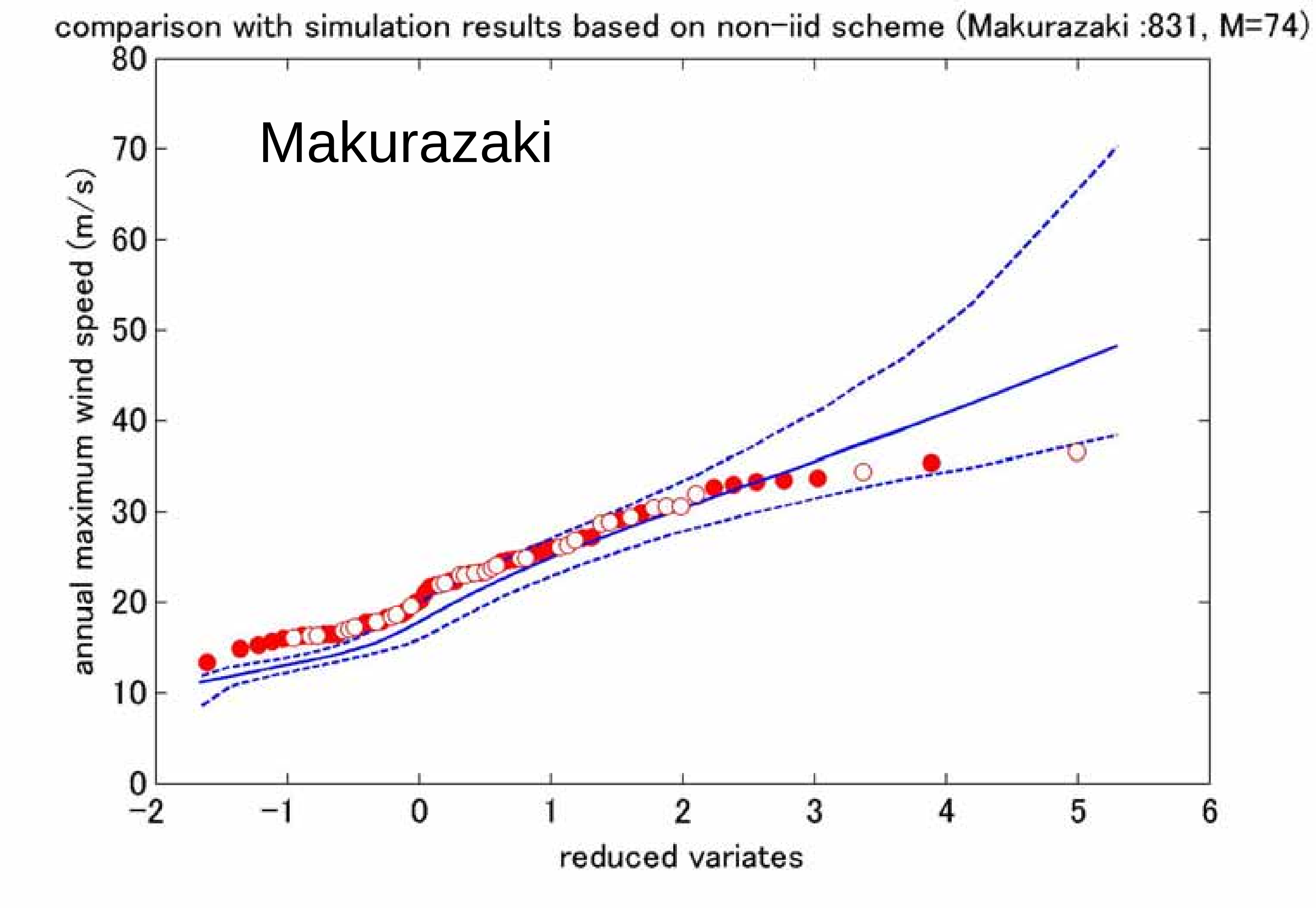
Sendai


comparison with simulation results based on non−iid scheme (Tokyo :662, M=74)
Tokyo
annual maximum wind speed (m/s)
reduced variates
40
35
30
25
20
15
10
−2
−1
0
1
2
3
4
5
6

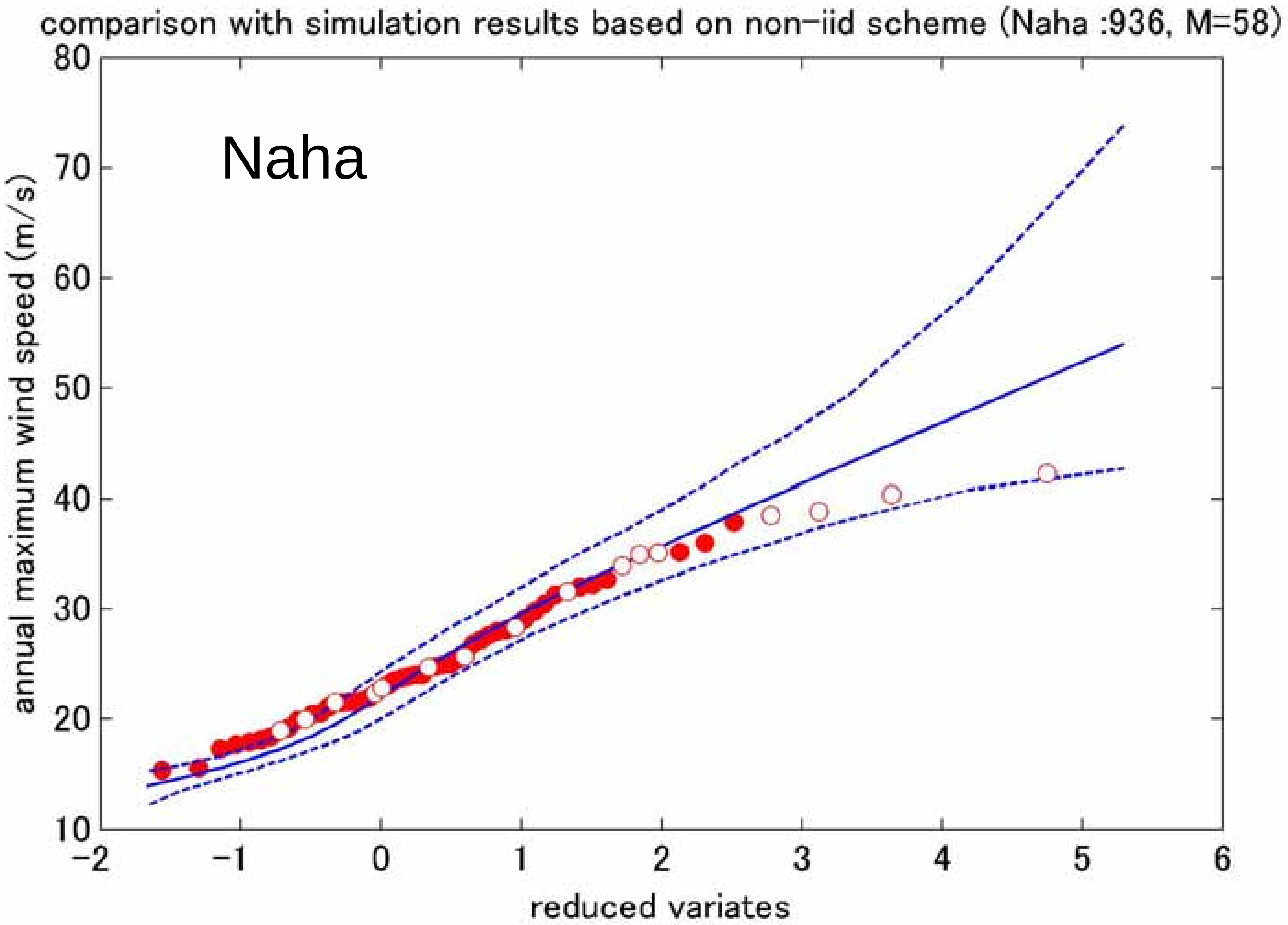

comparison with simulation results based on non-iid scheme (Kobe :770, M=74)
Kobe
annual maximum wind speed (m/s)
40
35
30
25
20
15
10
−2
−1
0
1
2
3
4
5
6
reduced variates


comparison with simulation results based on non-iid scheme (Shionomisaki :778, M=74)
Shionomisaki
annual maximum wind speed (m/s)
reduced variates
10
15
20
25
30
35
40
45
−2
−1
0
1
2
3
4
5
6

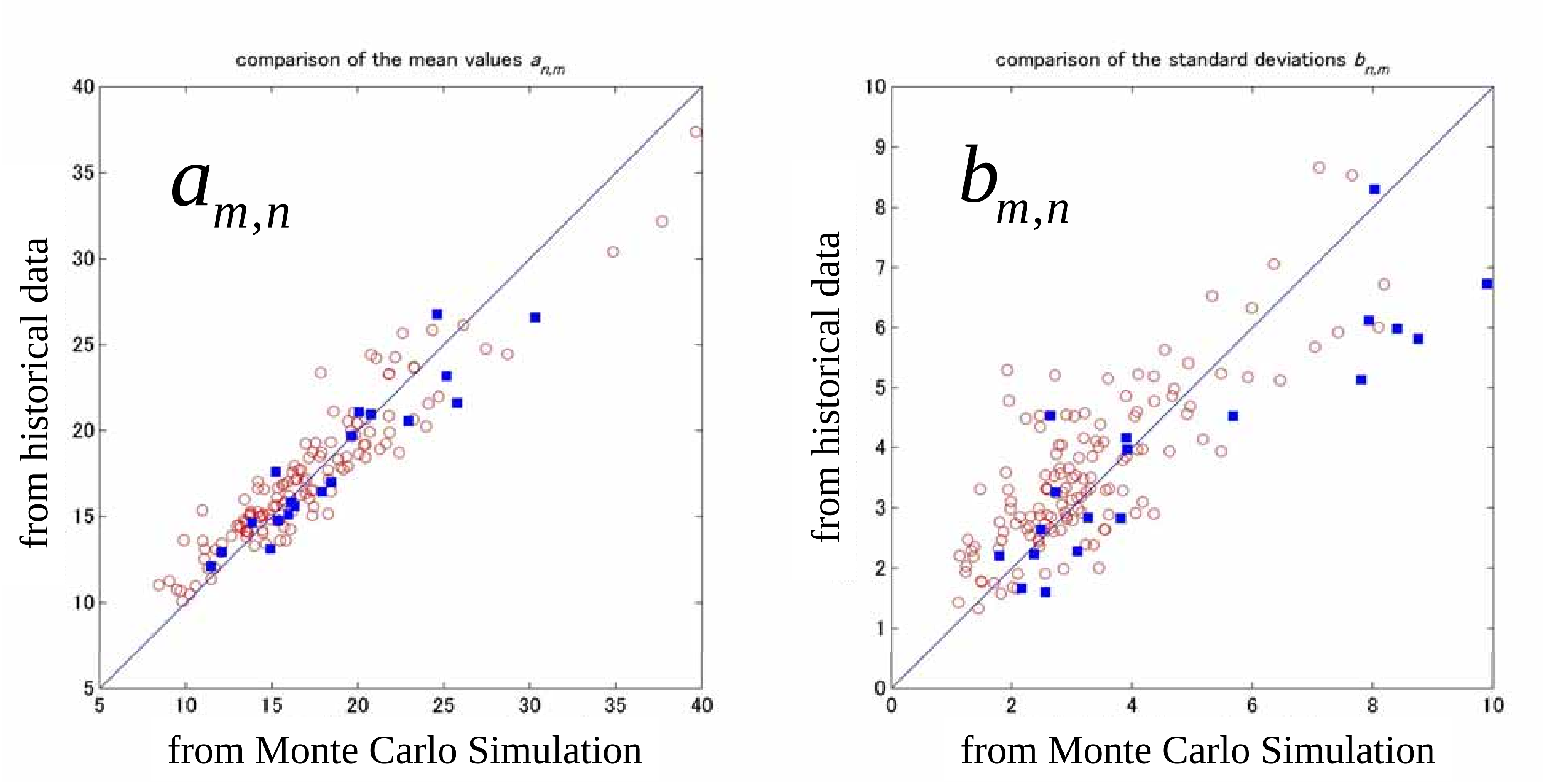
Makurazaki



comparison with simulation results based on non−iid scheme (Makurazaki :831, M=74)


comparison with simulation results based on non-iid scheme (Naha :936, M=58)
Naha
annual maximum wind speed (m/s)
80
70
60
50
40
30
20
10
−2
−1
0
1
2
3
4
5
6
reduced variates

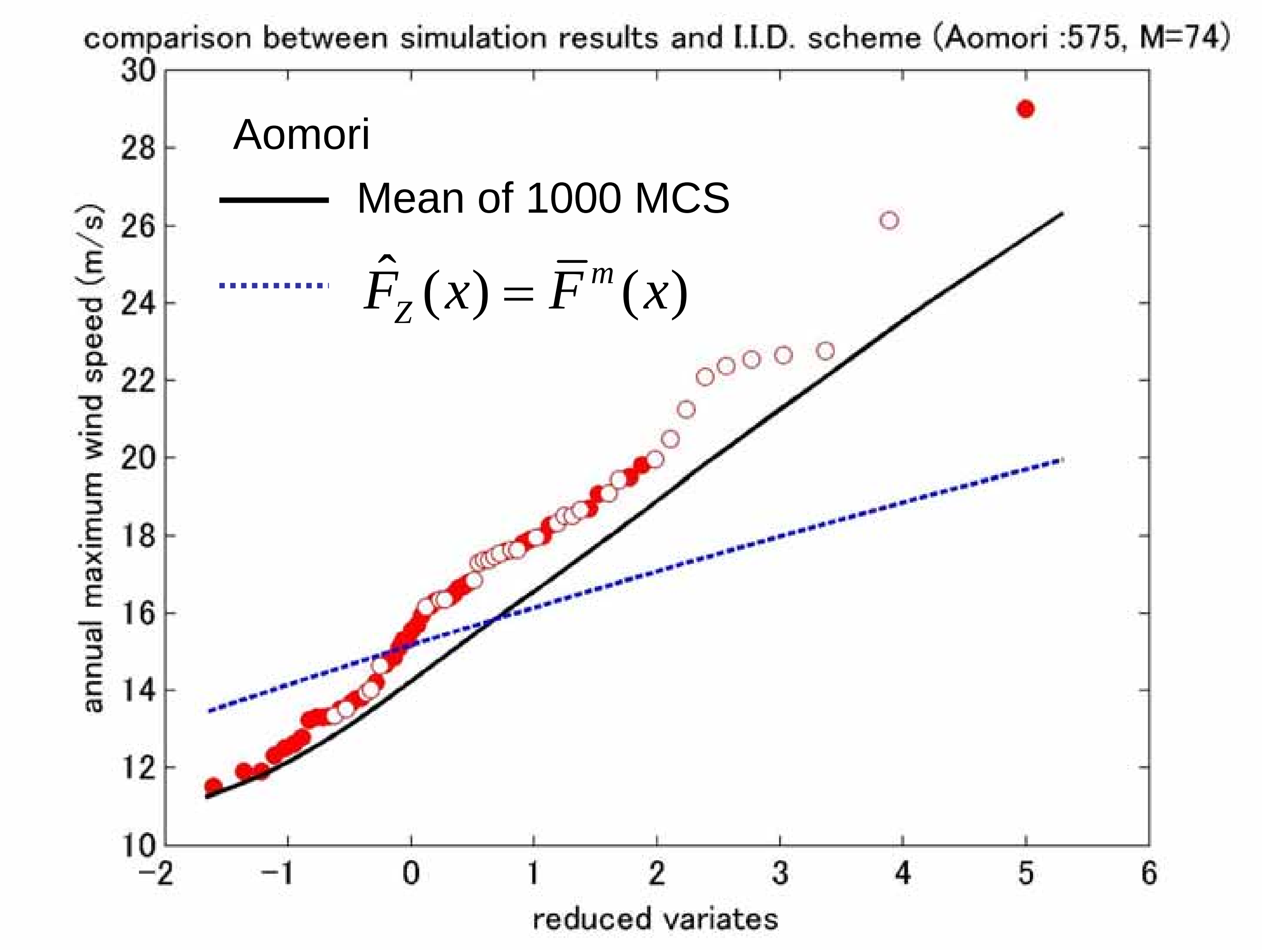
8) Comparison of the attraction coefficients from MCS and the historical records

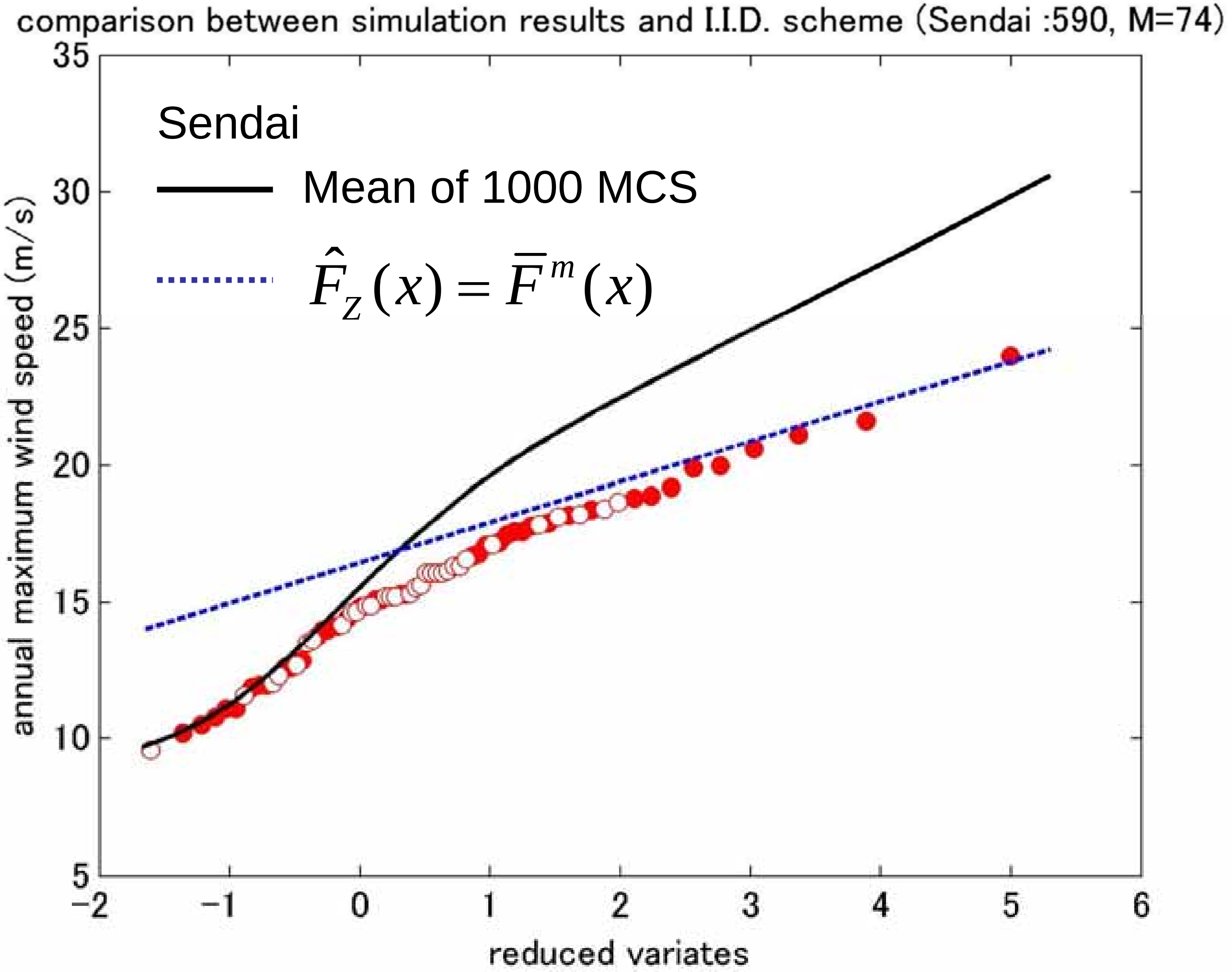


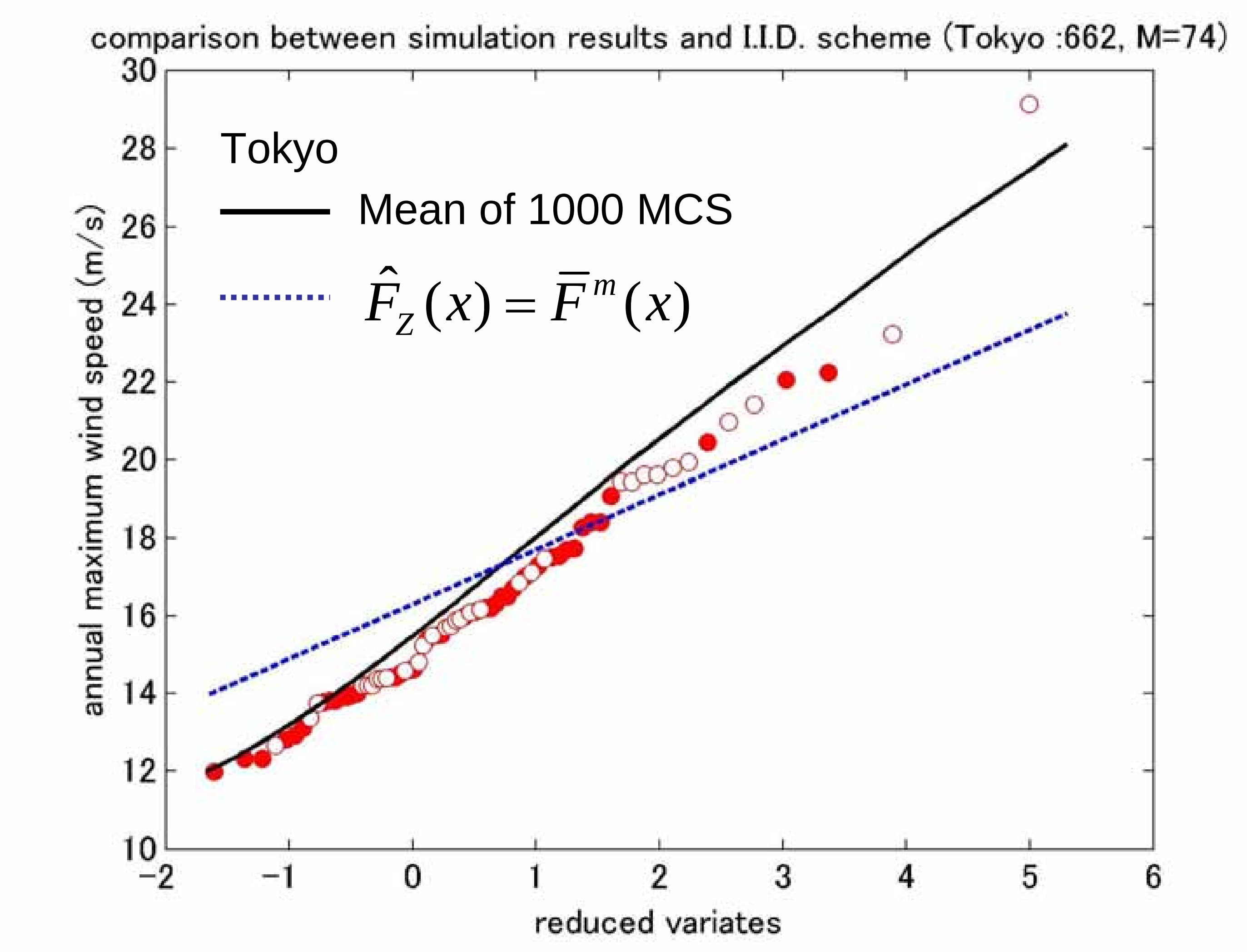
○ : n>50 (136 sites), ■ : n≦50 (19 sites)

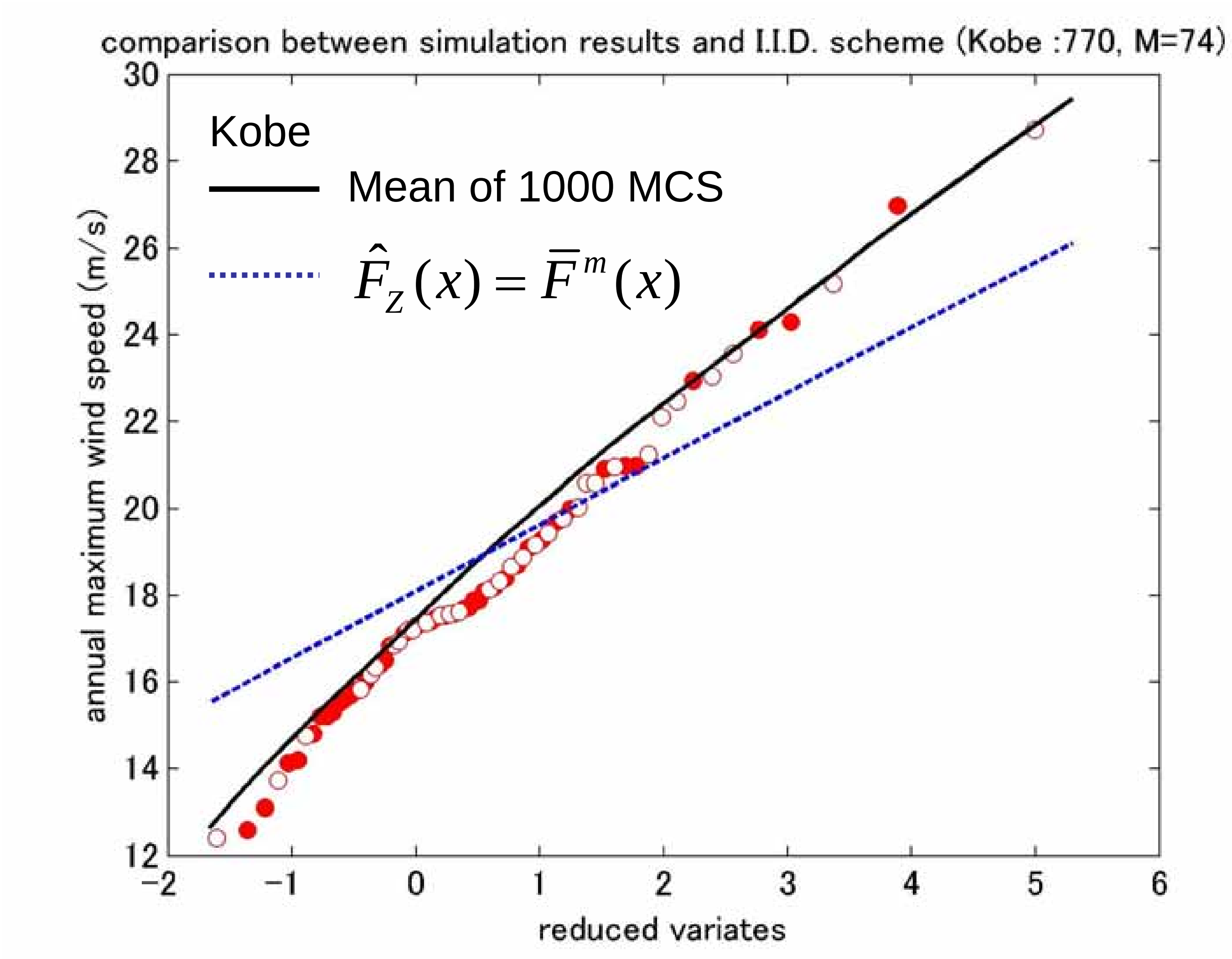
$$\lim_{n\to\infty} a_{m,n} = a_m \ , \ \lim_{n\to\infty} b_{m,n} = b_m$$

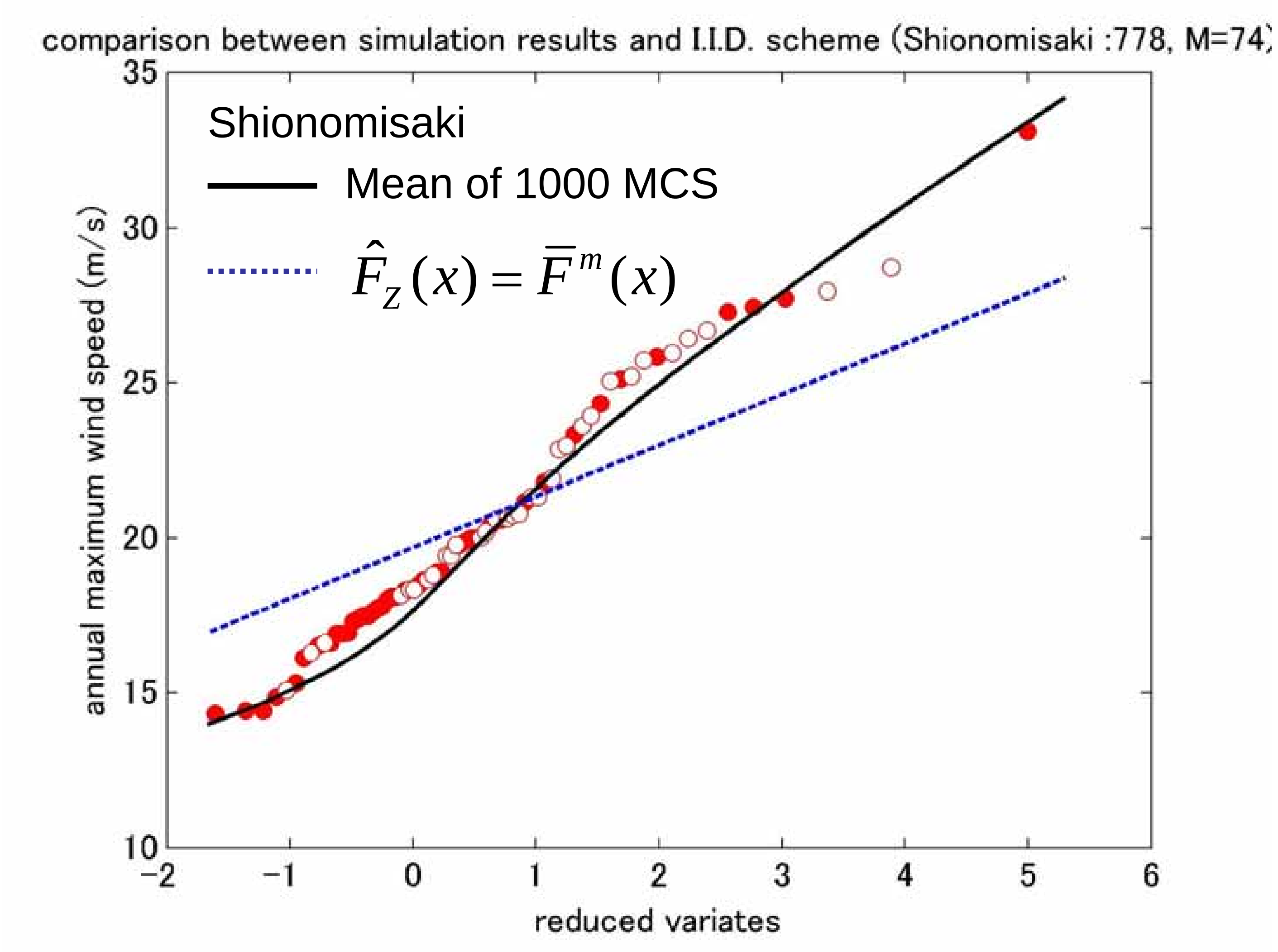
9) Defect of the alternative definition $\hat{F}_Z(x) = \bar{F}^m(x)$

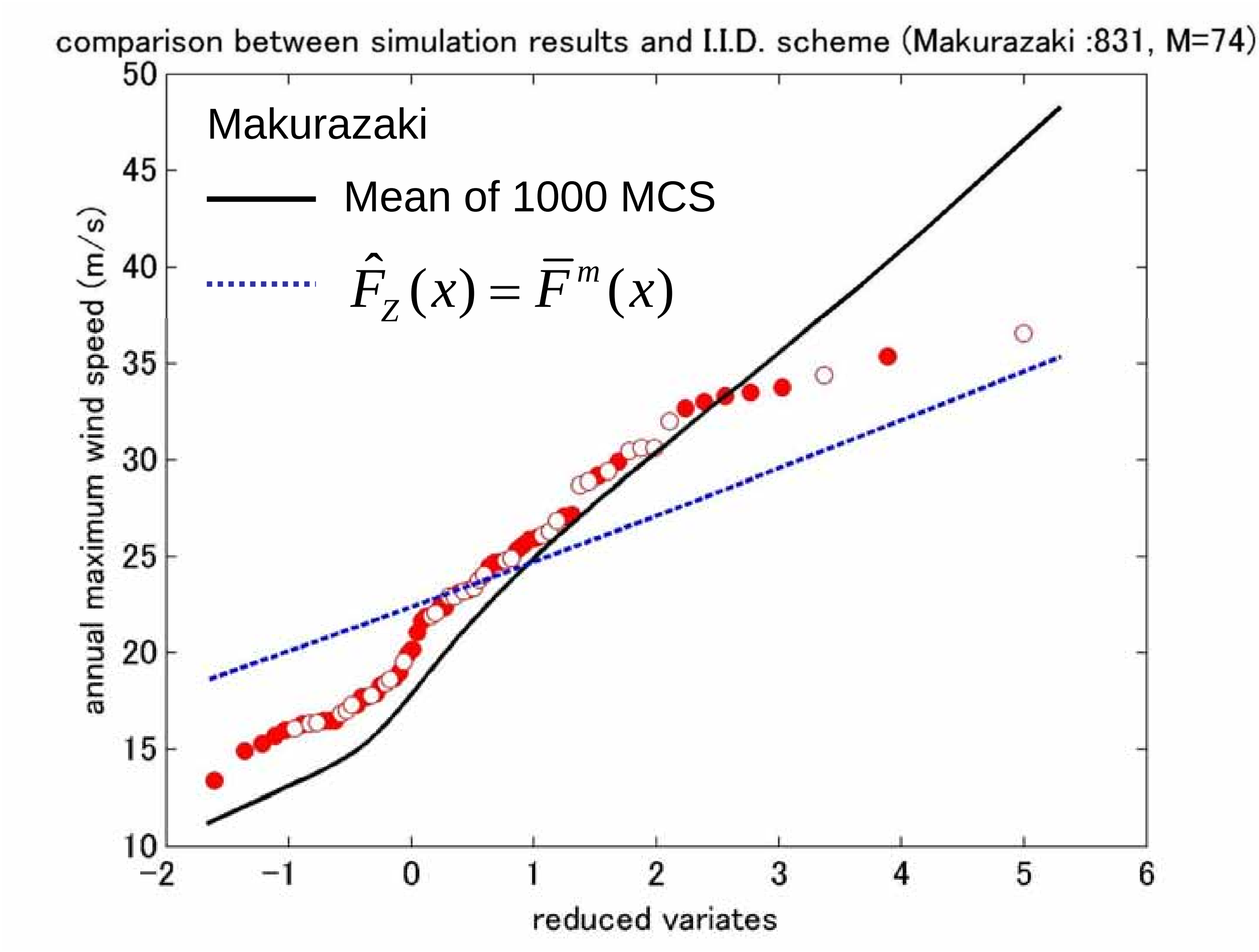


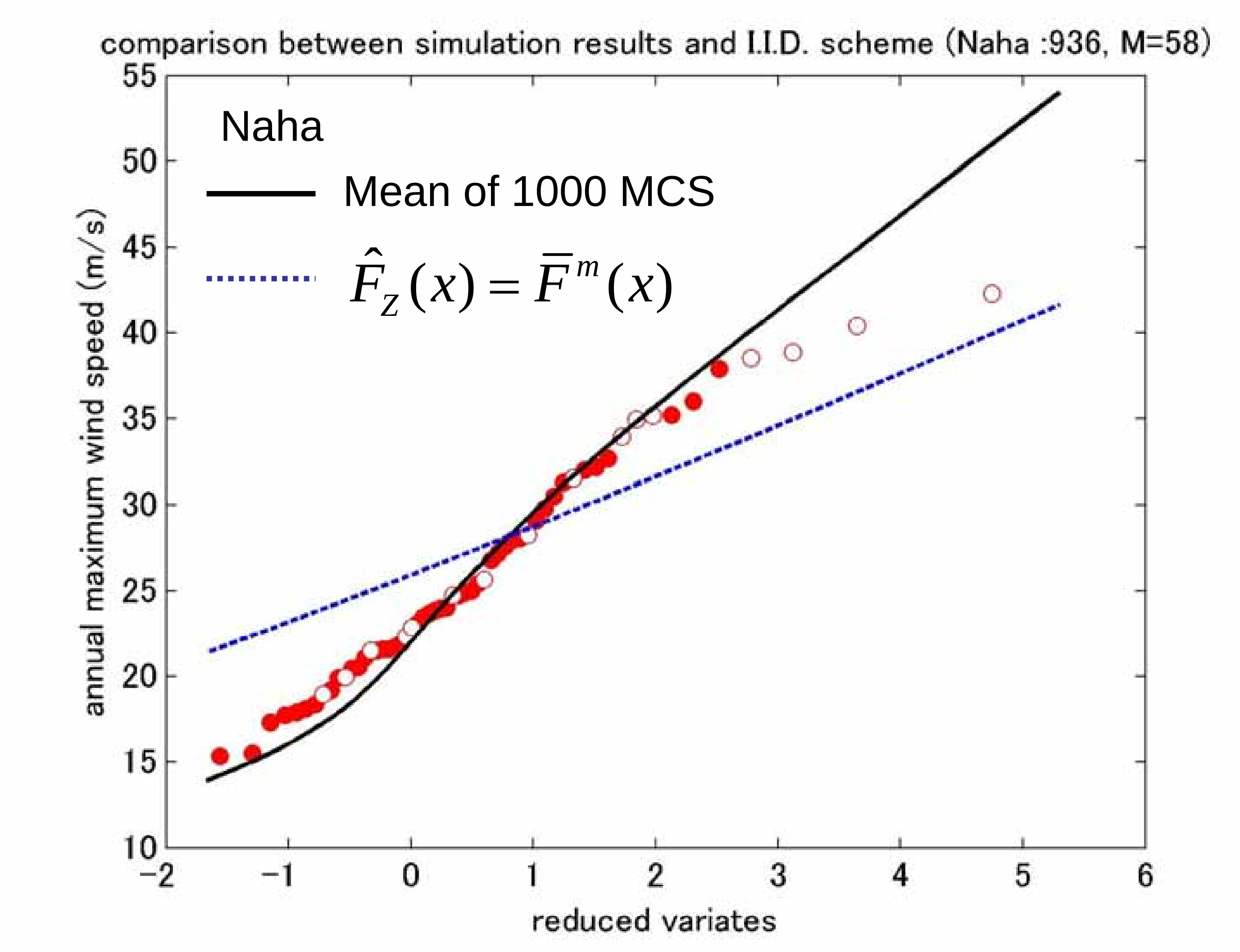

comparison between simulation results and I.I.D. scheme (Sendai :590, M=74)
Sendai
Mean of 1000 MCS
$\hat{F}_Z(x) = \overline{F}^m(x)$
annual maximum wind speed (m/s)
35
30
25
20
15
10
5
−2
−1
0
1
2
3
4
5
6
reduced variates


comparison between simulation results and I.I.D. scheme (Tokyo :662, M=74)
Tokyo
Mean of 1000 MCS
$\hat{F}_Z(x) = \bar{F}^m(x)$
annual maximum wind speed (m/s)
30
28
26
24
22
20
18
16
14
12
10
−2
−1
0
1
2
3
4
5
6
reduced variates

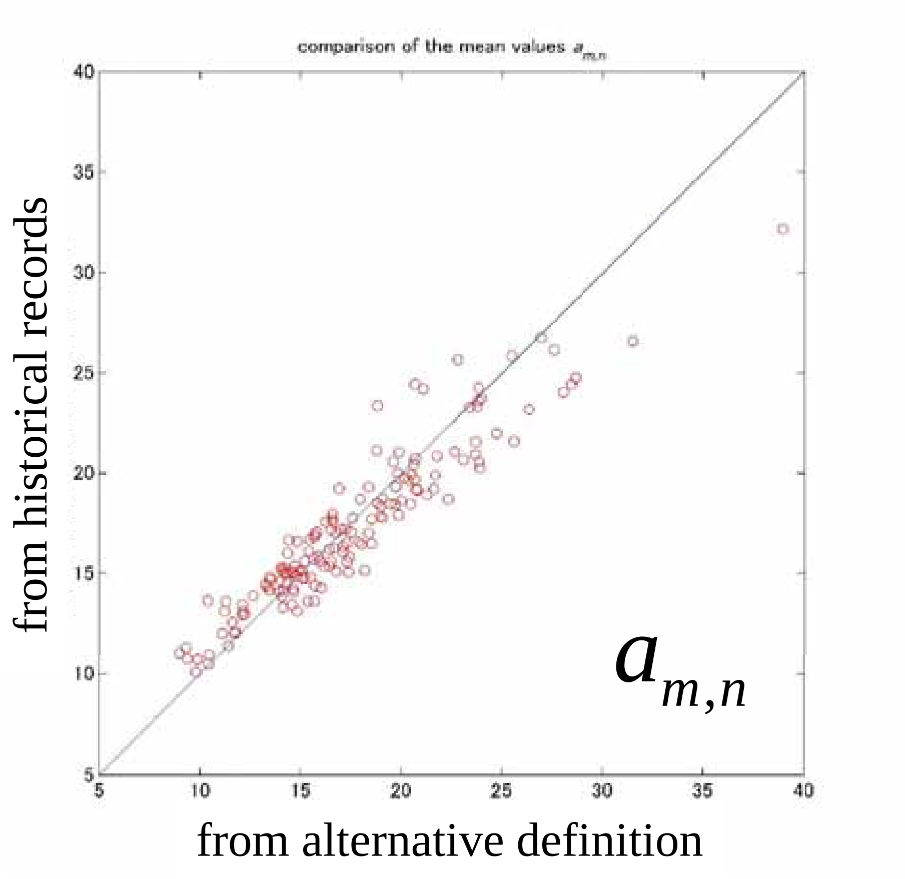

comparison between simulation results and I.I.D. scheme (Kobe :770, M=74)
Kobe
Mean of 1000 MCS
$\hat{F}_Z(x) = \overline{F}^m(x)$
annual maximum wind speed (m/s)
30
28
26
24
22
20
18
16
14
12
−2
−1
0
1
2
3
4
5
6
reduced variates

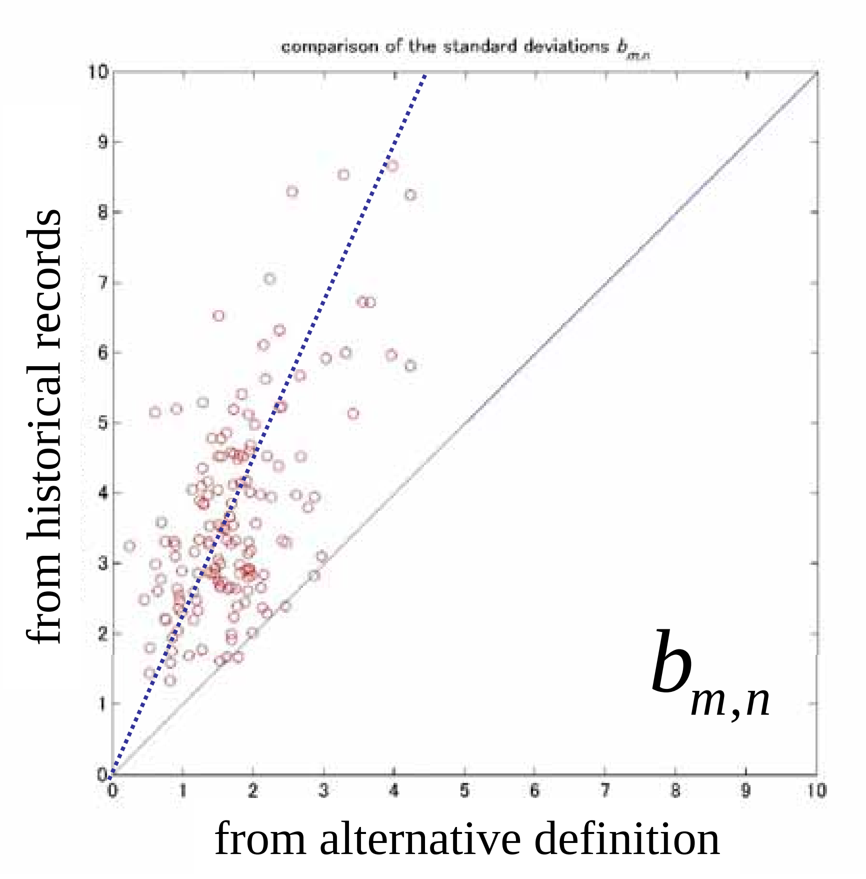

comparison between simulation results and I.I.D. scheme (Shionomisaki :778, M=74)
Shionomisaki
Mean of 1000 MCS
$\hat{F}_Z(x) = \bar{F}^m(x)$
annual maximum wind speed (m/s)
35
30
25
20
15
10
−2
−1
0
1
2
3
4
5
6
reduced variates


comparison between simulation results and I.I.D. scheme (Makurazaki :831, M=74)
Makurazaki
Mean of 1000 MCS
$\hat{F}_Z(x) = \bar{F}^m(x)$
annual maximum wind speed (m/s)
50
45
40
35
30
25
20
15
10
−2
−1
0
1
2
3
4
5
6
reduced variates


comparison between simulation results and I.I.D. scheme (Naha :936, M=58)
Naha
Mean of 1000 MCS
$\hat{F}_Z(x) = \bar{F}^m(x)$
annual maximum wind speed (m/s)
reduced variates
55
50
45
40
35
30
25
20
15
10
−2
−1
0
1
2
3
4
5
6

10) Comparison of the attraction coefficients from the alternative definition and the historical records

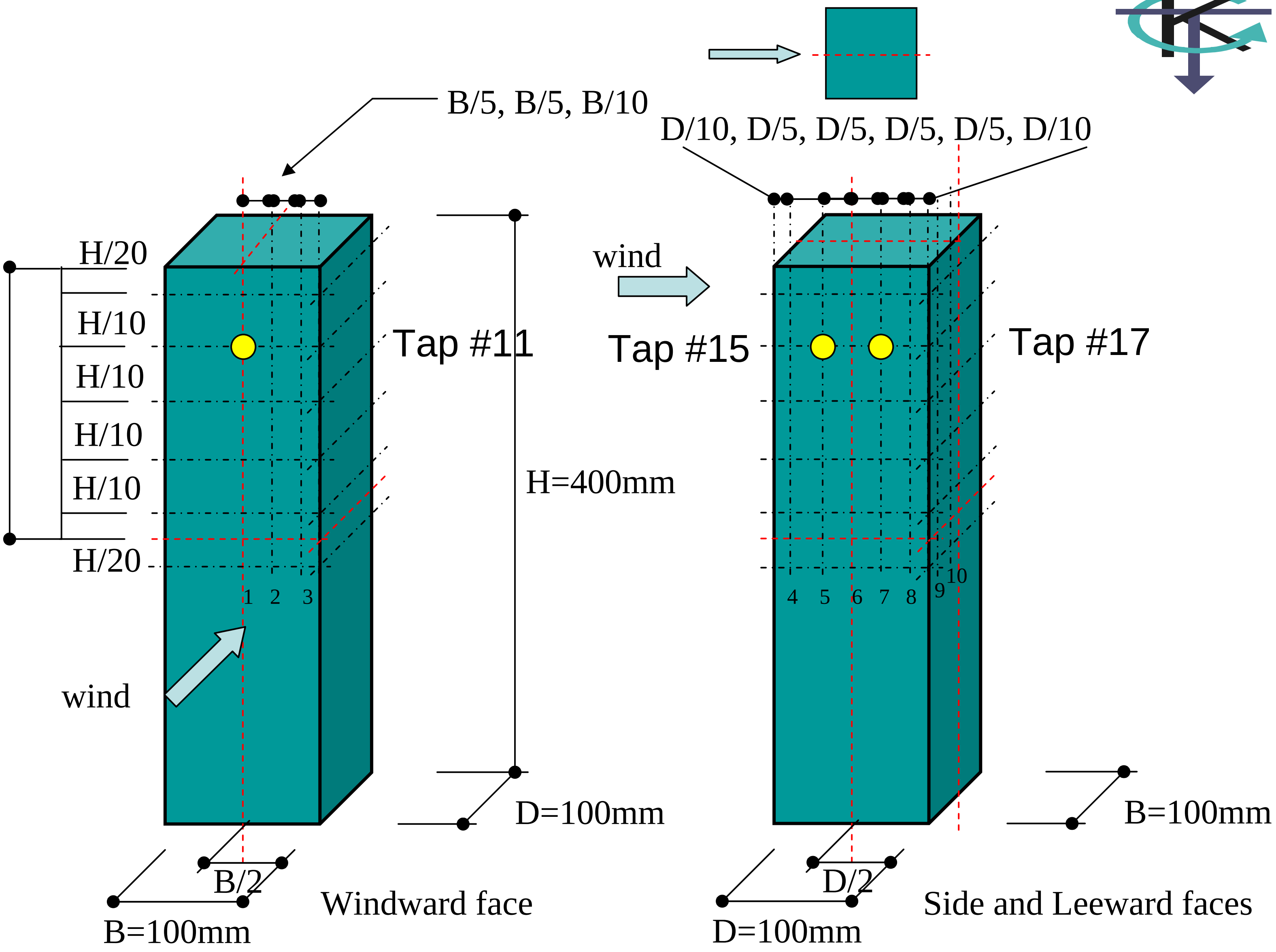




10 The cult of isolated statistics and The law of large number

How many extreme values should be used to estimate an extreme value distribution?
(case study for max/min pressure coefficients)*

* Choi & Kanda (2004), Stability of extreme quantile function estimation from relatively short records having different parent distributions, *Proc. 18th Natl. Symp. Wind Engr*., p455~460 (in Japanese)

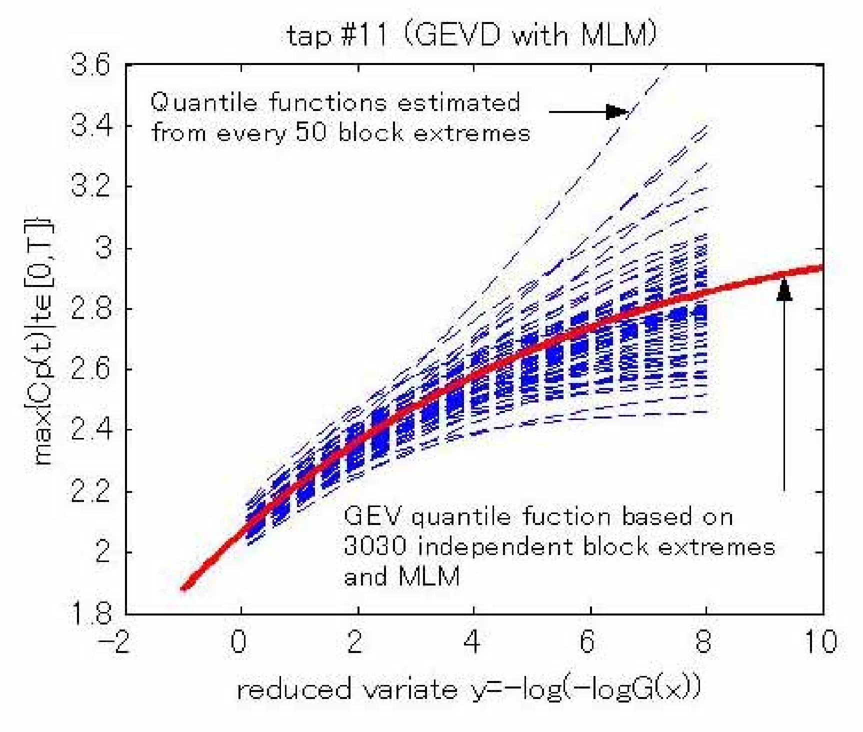

B/5, B/5, B/10
D/10, D/5, D/5, D/5, D/5, D/10
H/20
H/10
H/10
H/10
H/10
H/20
Tap #11
Tap #15
Tap #17
wind
wind
H=400mm
1
2
3
4
5
6
7
8
9
10
D=100mm
B=100mm
B/2
D/2
B=100mm
D=100mm
Windward face
Side and Leeward faces

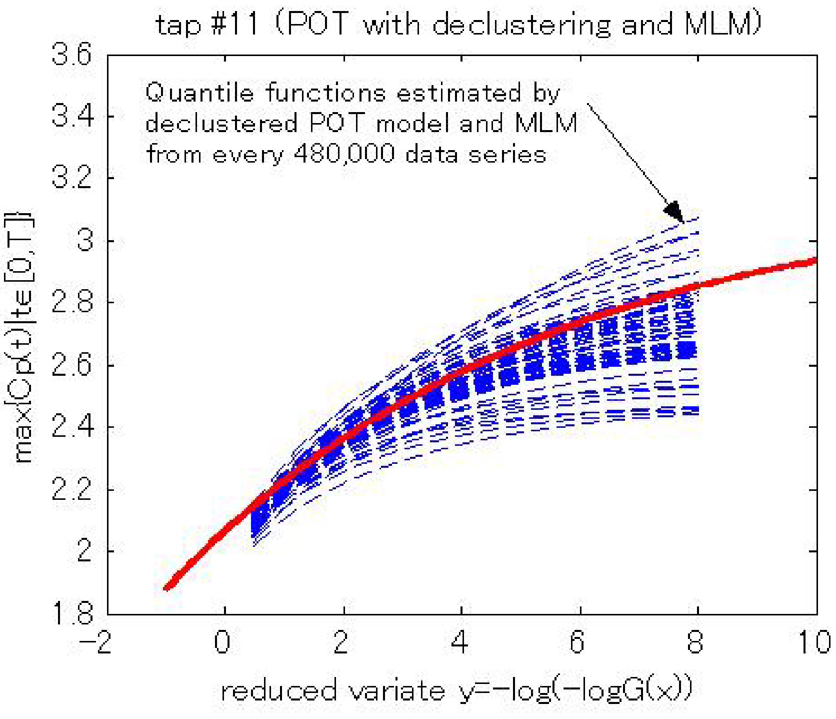
Which one is the best estimation?

$$C_p(t) = \frac{p(t) - p_s}{1/2\rho U_H^2}$$

$p(t)$: instant total pressure, p_s : static pressure

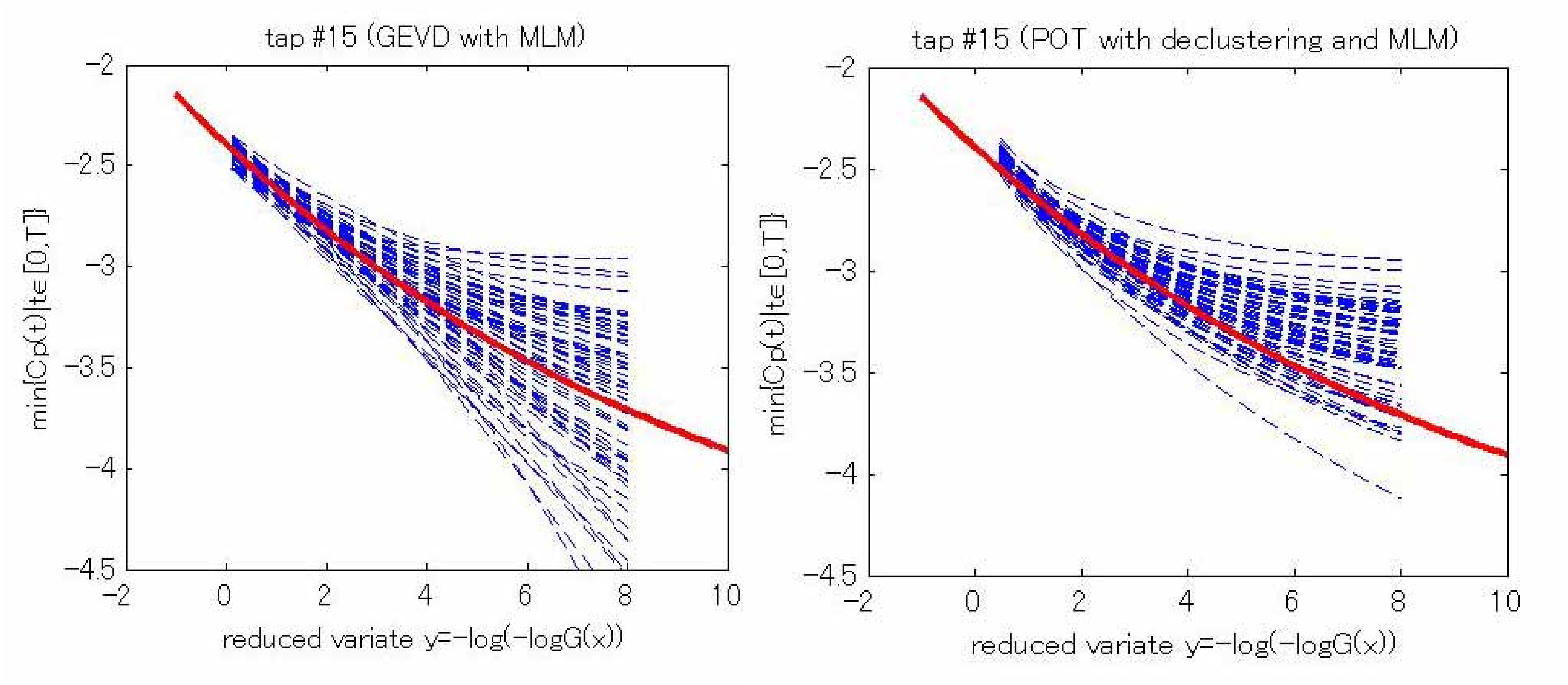
ρ : air density, U_H : wind velocity at model height





* 50 blocks contain about 480,000 discrete data

Which one is the best estimation?



10 The cult of isolated statistics and The law of large number

We never be free from the law of large number.

The statistician and the scientists/technologist need to understand that models are necessarily simplifications of the system being modelled; that they are , an absolute sense, wrong; that they are certainly provisional, but nonetheless are useful and necessary for successful quantitative thinking.

from J.A. Nelder (1986), Statistics, Science and Technology – *The address of the president, delivered to the* Royal Statistical Society *on Wednesday, April 16th, 1986, J. R. Statist. Soc. A* 149(2), p109~121