

## What is uncertainty of the estimation?

Vladimir Sokolov and Friedemann Wenzel, Geophysical Institute of Karlsruhe University, Karlsruhe, GERMANY

### MAGNITUDE

#### Summary

Determination of earthquake magnitude and strength of shaking at a site from the initial P-wave portion of ground motion is the key problems for earthquake Early Warning (EEW). We analyzed performance of scaling relations between so-called characteristic period  $\tau_c$  and moment magnitude  $M_w$  in respect of (a) characteristics of datasets accumulated in various regions (earthquake depth, characteristics of network) and (b) variation of initial conditions applying for determination of the parameter  $\tau_c$  (length of P-wave windows, number of used stations). The used data contain strong-motion records from 110 earthquakes (moment magnitude range 4.4 - 7.6) occurred in Japan and Taiwan.

We show that, although the standard error of regression  $\tau_c = f(M_w)$  in general becomes smaller with the increase of length of P-wave window (PL) and number (N) of averaged observations (stations), the uncertainty in estimation of given observed does not decrease further for  $PL > 3-4$  sec and  $N > 3-4$  stations. The information about earthquake depth plays an important role in reducing uncertainty of magnitude estimations.

**Magnitude of an earthquake may be rapidly estimated using the frequency content of the first few seconds after P-wave arrival.**

As has been found from several studies, there is a stable parameter for the magnitude determination, so-called "characteristic period" (Kanamori 2005).

$$\tau_c = 2\pi/\sqrt{r} \quad r = \int_0^{\tau_c} \dot{u}^2(t) dt / \int_0^{\tau_c} u^2(t) dt$$

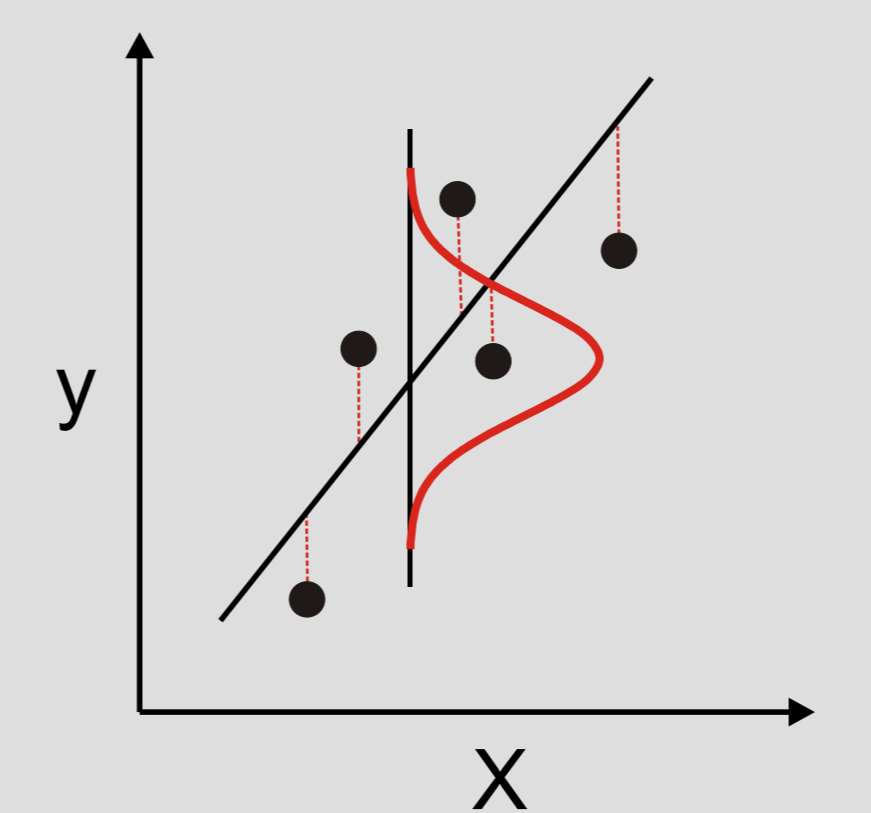
The vertical component of ground motion is used and the high-pass filtering with a cut-off frequency of 0.075 Hz is applied to remove low-frequency noise.

The regression analysis is applied to determine the dependence between selected parameter,  $\tau_c$  (Y, response) and magnitude (X, predictor)  $\tau_c = f(M_w)$

In **ordinary least squares (OLS)**, we should assume that the independent variables (X, magnitude) are measured without error and all of the errors are in the dependent (response) parameter variables (Y, characteristic period). The errors are characterized by **normal distribution** and **zero mean** (see Figure to the right).

In OLS the regression line should be fit by minimizing the squared **vertical distances** between the data and the line

#### OLS regression



### INTENSITY

#### Summary

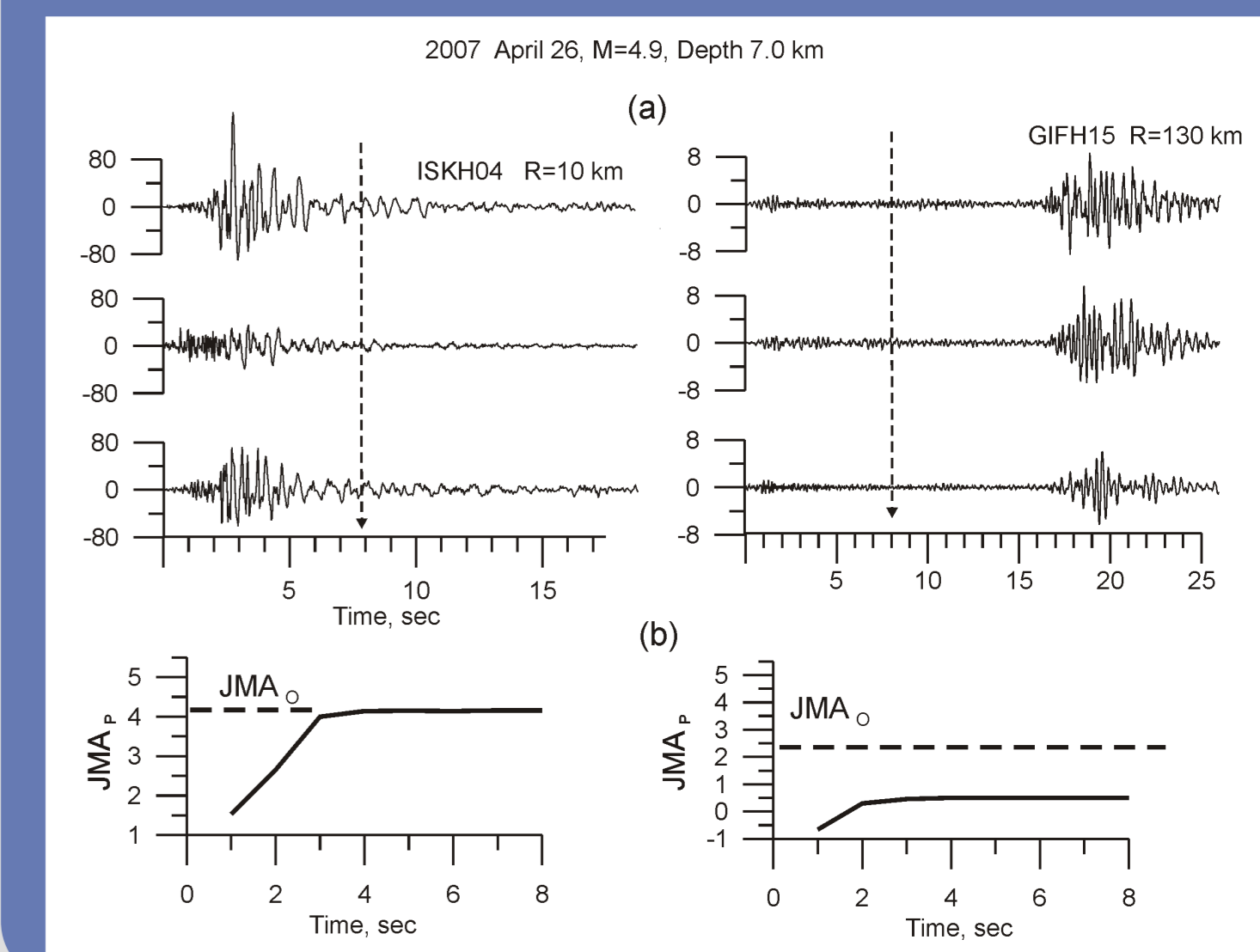
The most of the techniques proposed for estimation of intensity of shaking utilize (a) ground motion models based on the estimated magnitude and hypocentral distance, or (b) the interim proxies, such as initial vertical displacement or so-called Damage Index.

We suggest the instrumental Japan Meteorological Agency intensity ( $JMA_i$ ) as a characteristic for fast estimation of damage potential in the EEW systems. We investigated the scaling relations between  $JMA_i$  measured using the whole earthquake recordings (overall intensity) and using particular time intervals of various duration (1.0 - 8.0 seconds) starting from the P-wave arrival (preliminary intensity). The dataset included 3660 records (K-NET and the KiK-net networks) from 55 events ( $M_w$  4.1 - 7.4) occurred in 1999-2008 in Japan. We showed that the time interval of 4-5 seconds from the P-wave arrival can be used for reliable estimations of the overall intensity with the average standard error of about **0.5 JMA units**. The uncertainty in the prediction may be reduced by the consideration of local site conditions or by the development of the station-specific models.

#### Estimation of JMA intensity for earthquake early warning purpose.

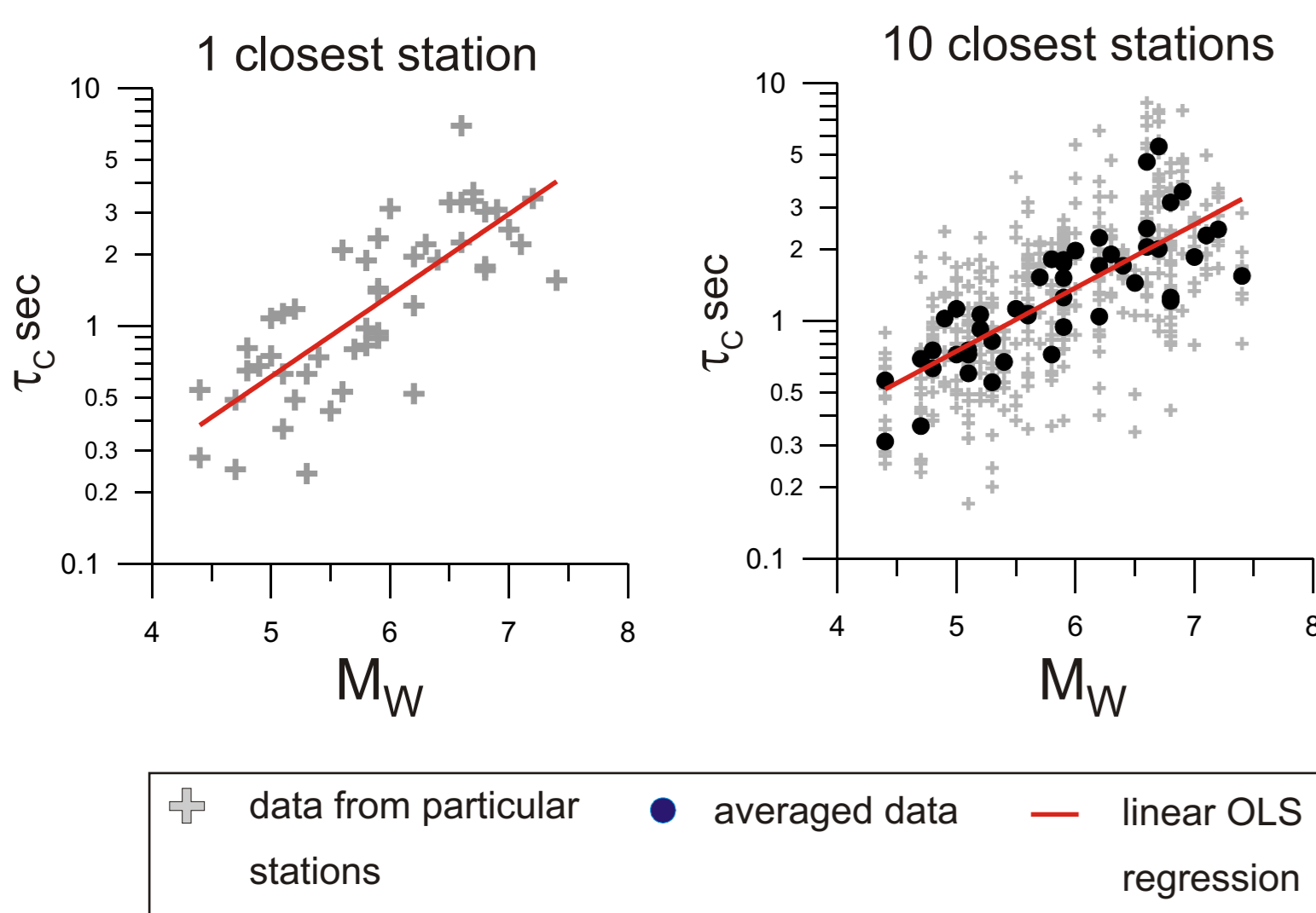
(a) Ground motion records (acceleration,  $cm/s^2$ ) obtained at two stations located at different hypocentral distances R from the same earthquake. Dashed lines mark the first 8 seconds from the P-wave arrival.

(b) Time-dependent nature of the  $JMA_p$  intensity estimated from the initial portions of P-wave using various time intervals. Dashed line show the values of instrumental intensity calculated for the whole earthquake record ( $JMA_o$ ).



### DIRECT relationship $\tau_c = f(M_w)$

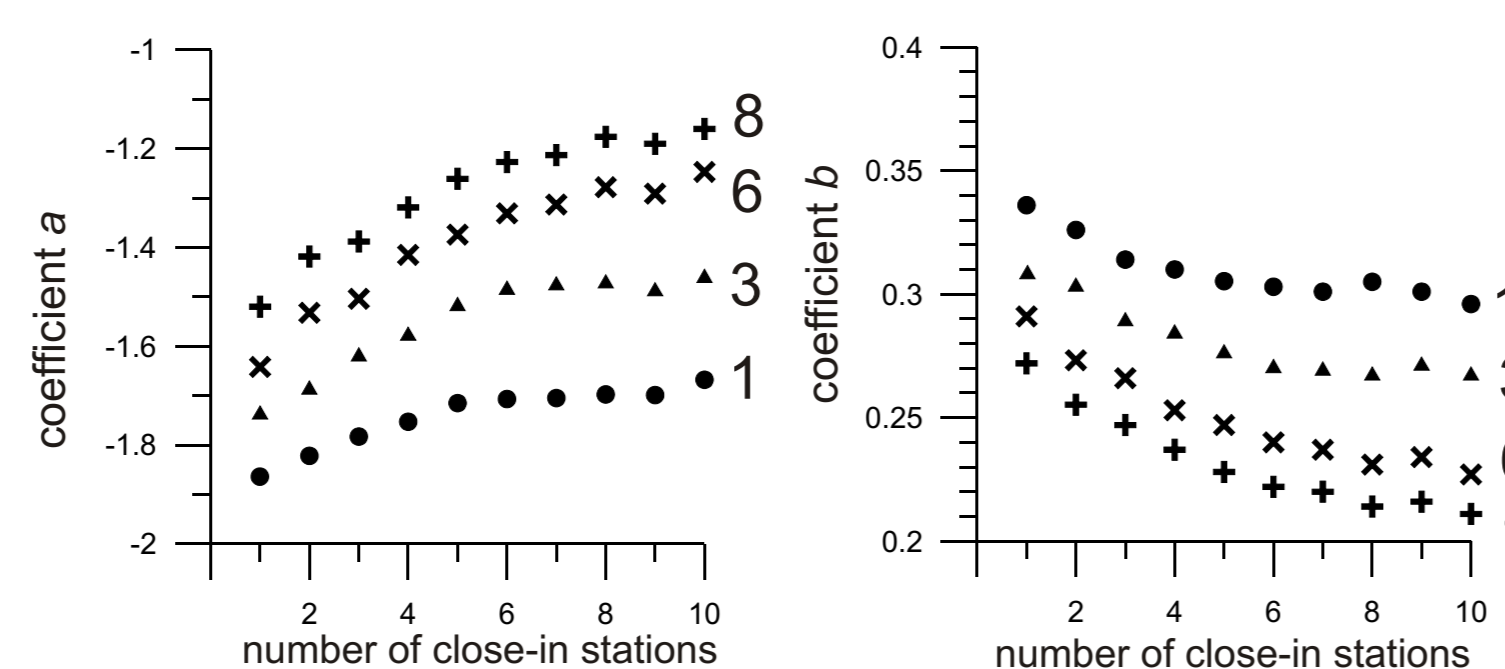
#### Examples of linear relationship $\tau_c = f(M_w)$



#### Distribution of coefficients a and b in relationship

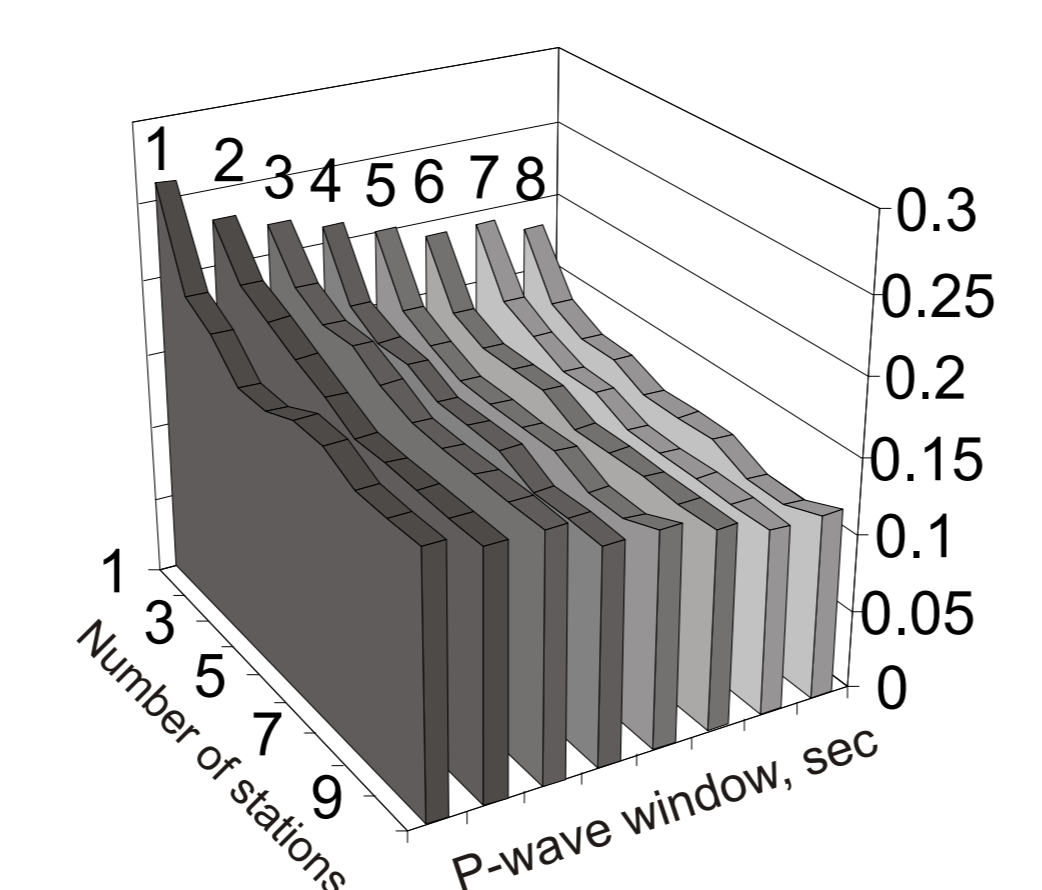
$$\log_{10} \tau_c = a + bM_w \pm \sigma$$

for various numbers of used stations and duration of P-wave window (1, 3, 6, and 8 sec)

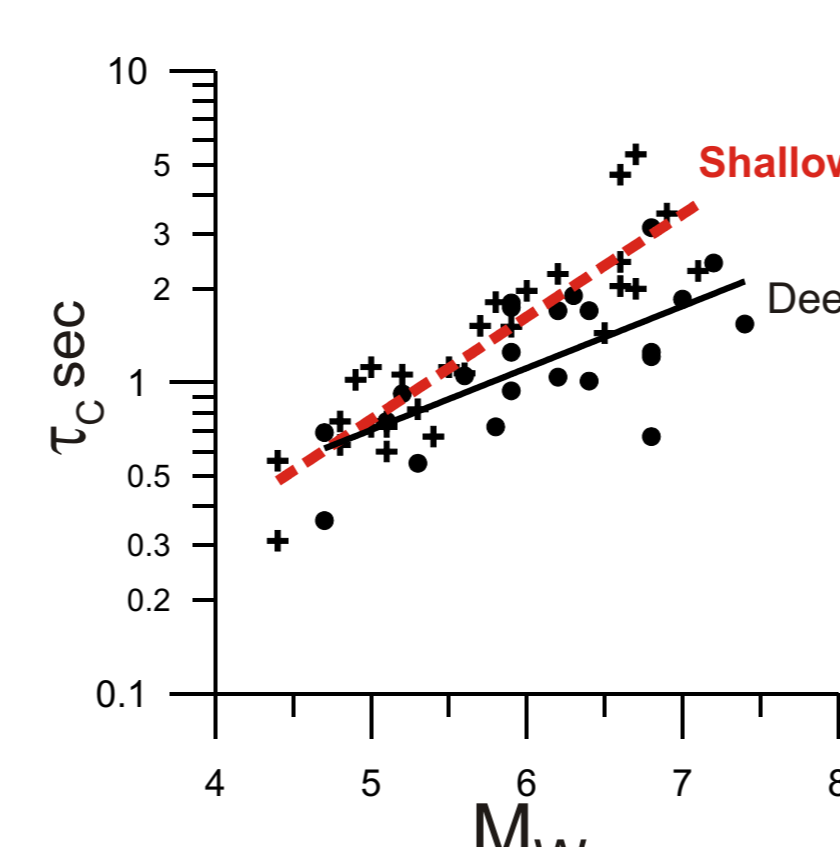


#### Distribution of standard errors

$\sigma$  of the relationship  $\tau_c = f(M_w)$

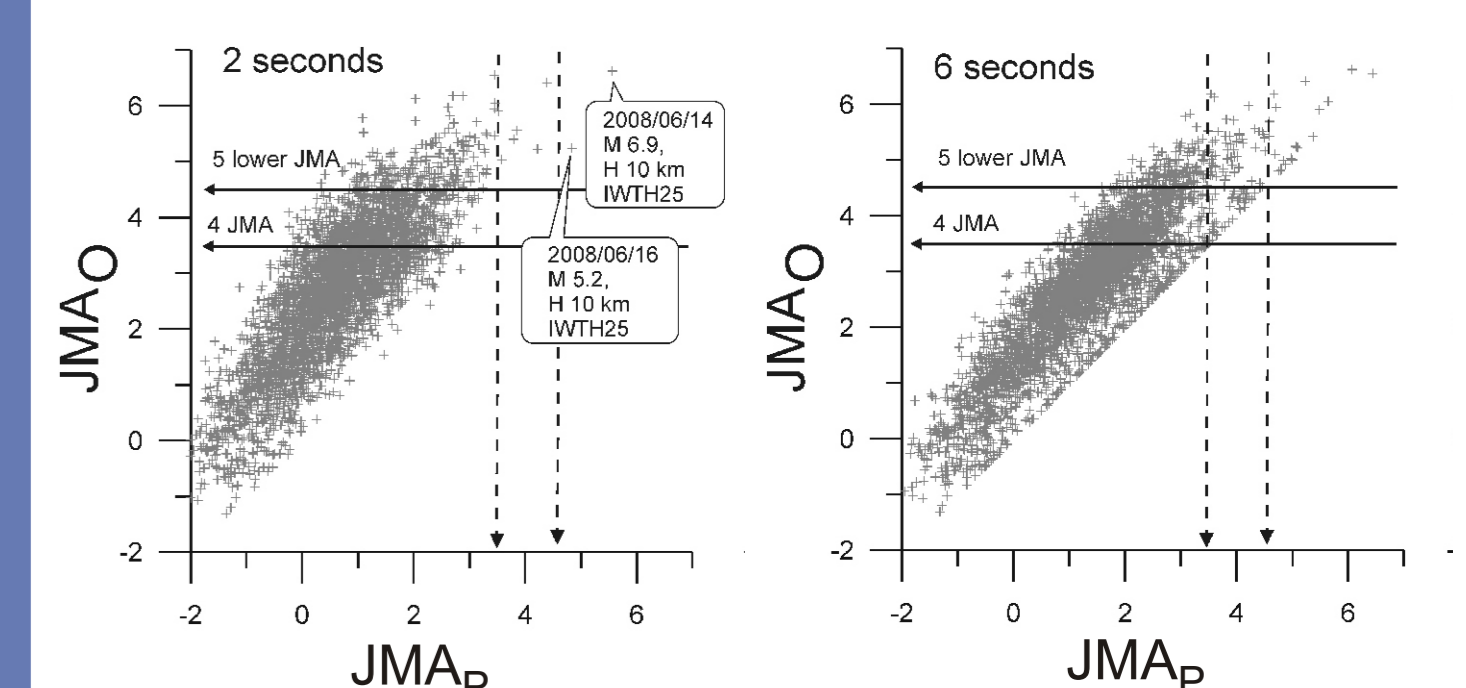


#### Relationships $\tau_c = f(M_w)$ for shallow and deep earthquakes in Japan



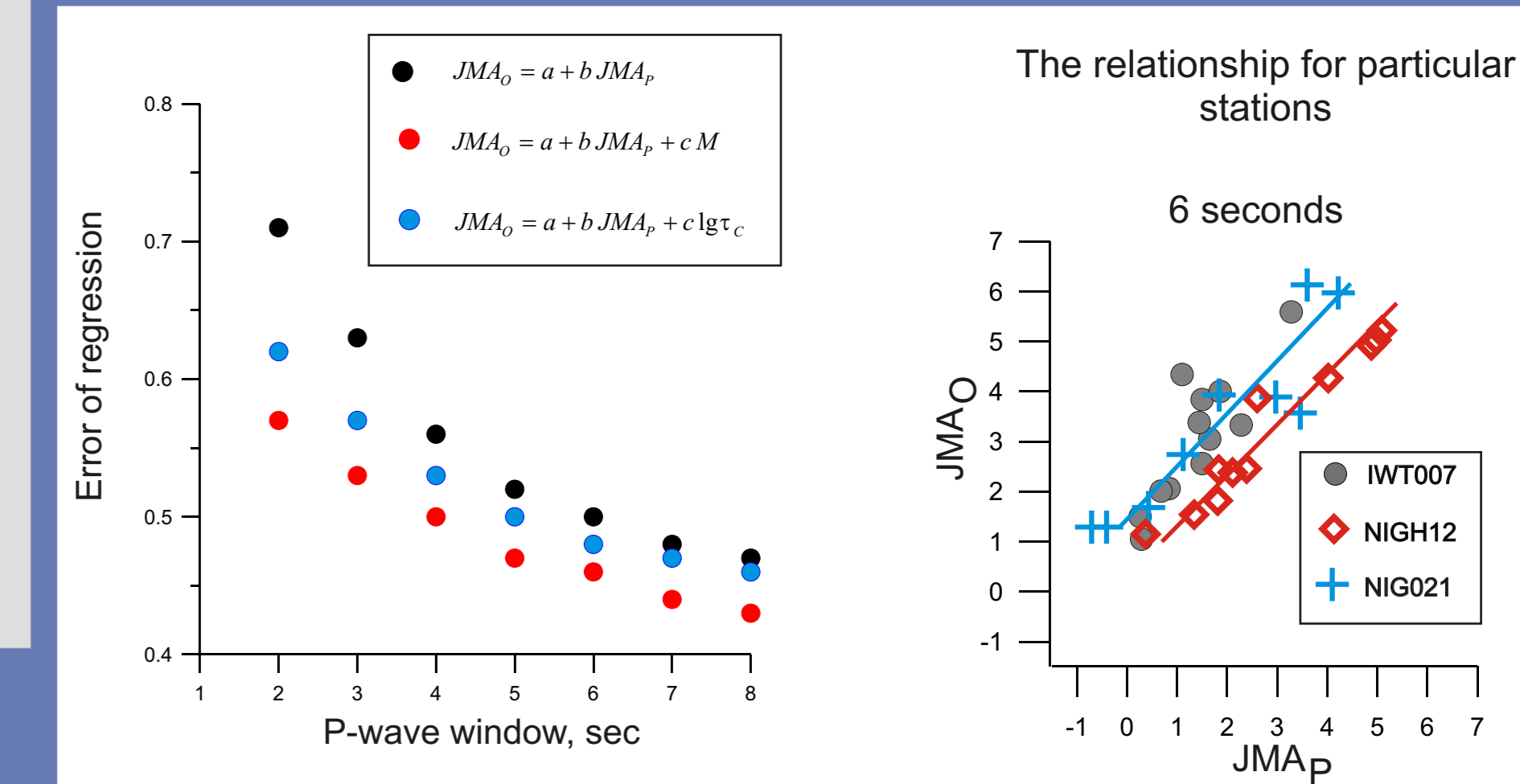
#### In Japan the EEW issues consider two levels of JMA intensities, namely: 4 and 5 lower (5-).

Distribution of the JMA instrumental intensity calculated for the whole earthquake records ( $JMA_o$ ) versus the intensity estimated for the initial portions of ground motion ( $JMA_p$ ) using various time intervals is shown on Figure to the right.



#### We considered three types of relationship between overall intensity ( $JMA_o$ ) and preliminary intensity ( $JMA_p$ ) using simple relation $JMA_o = a + bJMA_p$ and also Magnitude $JMA_o = a + bJMA_p + cM$ and characteristic period $JMA_o = a + bJMA_p + c \lg \tau_c$

As can be seen from the picture, the standard error of regression becomes smaller, when increasing the P-wave window and considering magnitude of earthquake or  $\tau_c$ .



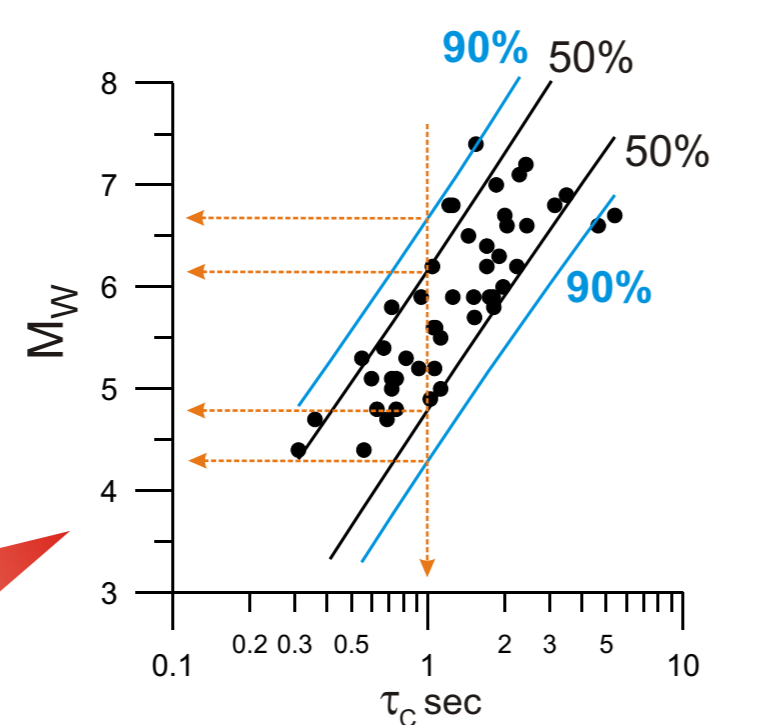
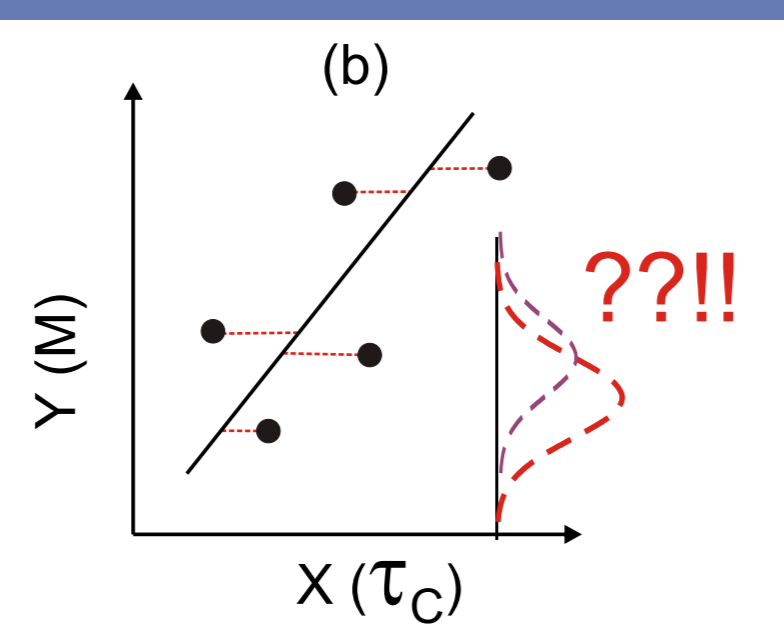
### REVERSE relationship $M_w = f(\tau_c)$

#### Now, we would like to estimate the true value M, given an observed $\tau_c$ .

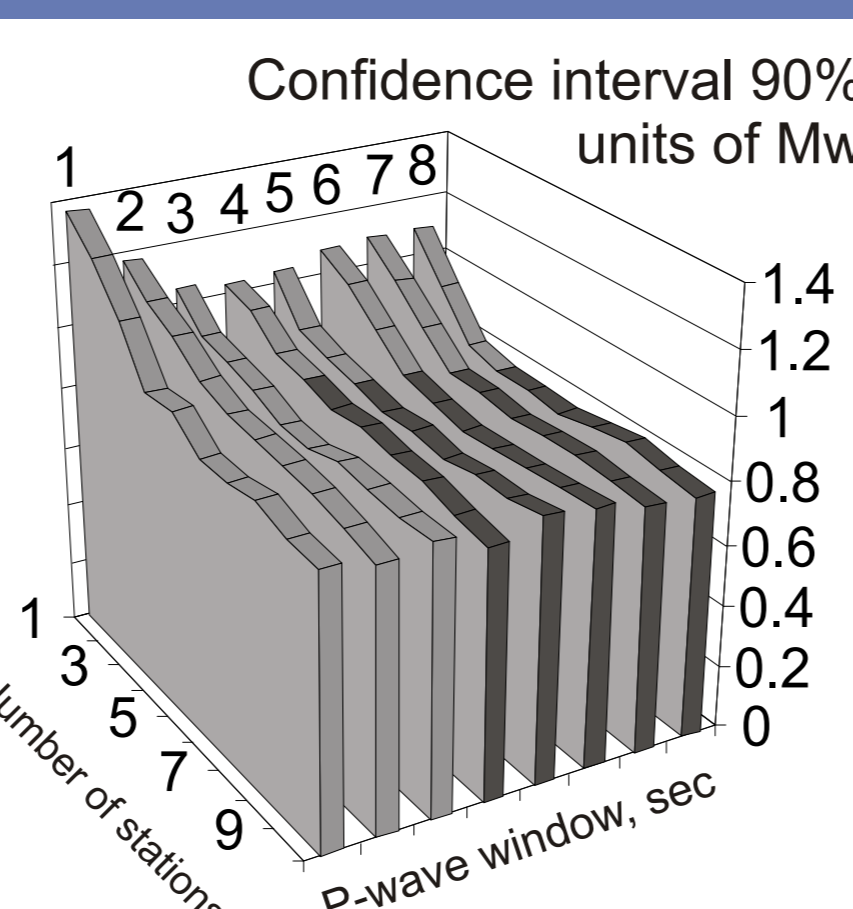
The estimation is equivalent to application of so-called **inverse (or reverse) regression** technique. In this case the variable to be predicted (M) is measured with negligible errors, as compared with errors of independent variable  $\tau_c$ . The regression line should now be fit by minimizing the squared **horizontal distances** between the data and the line (see Figure to the right)

In this case, **the estimator  $M_w = f(\tau_c)$  CANNOT** be considered as a mean of possible M values. The OLS calculation of **regression residuals**, i.e. the VERTICAL departures of each point M from the line, **could not be used** for assessment of uncertainty.

**The confidence limits for M for a given  $\tau_c$  should be used instead of single M value**



The **confidence limits** depend on number of observations N; coefficients b and standard error  $\sigma$  in direct regression  $\tau_c = f(M_w)$ . The smaller b – the larger width of confidence intervals. Thus, even if standard error may be reduced, the confidence intervals in **reverse regression** may be still wide.



Distribution of confidence limits (half-width) of the reverse relationship  $M_w = f(\tau_c)$

#### The scaling relationship between the characteristic period $\tau_c$ and moment magnitude $M_w$ allows predicting earthquake magnitude with accuracy of $\pm 0.8$ units for 90% and $\pm 0.45$ units for 59% confidence limits, when

- (1) using time interval of at least 3-4 seconds from P-wave arrival
- (2) averaging the data from at least 3-4 close-in stations.

**Every attempt to reduce the errors in the direct relationship  $\tau_c = f(M_w)$  should be accompanied by analysis of the confidence intervals for correspondent inverse estimations  $M_w = f(\tau_c)$**

## CONCLUSION

**The time interval of 4-5 seconds from the P-wave arrival may be considered as sufficient for reliable estimations of intensity of the strongest part of shaking with average standard error of about 0.5 units of JMA.**

**Empirically-derived station-dependent coefficients may reduce uncertainty.**