THE ROLE OF OCEAN STATE INDICES IN SEASONAL AND INTER-ANNUAL CLIMATE VARIABILITY OF THAILAND

Werapol Bejranonda\textsuperscript{1} and Manfred Koch\textsuperscript{2}

Department of Geohydraulics and Engineering Hydrology
University of Kassel, Kassel, 34190, Germany
E-mail: \textsuperscript{1}bwerapol@uni-kassel.de; \textsuperscript{2}kochm@uni-kassel.de

ABSTRACT
Thailand has a long coastline with the Pacific Ocean as well as with the Indian Ocean. Because of this peculiar location, Thailand’s local climate and, consequently, its surface water resources are strongly influenced by the mix of tropical dry and tropical wet (monsoon) seasons which, in turn, depend themselves upon the thermal states of the Pacific and Indian Oceans. Because of this teleconnective oceanic impact on local climate variations in different regions of Thailand, which acts not only on the seasonal but the inter-annual time scale as well, the somewhat extreme climate pattern, such as long periods of drought in some provinces of Thailand, must be seen in this light. With a better understanding of these ocean-Thai seasonal-weather-pattern teleconnections, water authorities in Thailand would get a tool at hand to forecast short-term extreme seasonal climate pattern in a particular region allowing them to better manage the adequate supply of surface water resources to the users, especially, rice farmers who support a large portion of the economy of the country.

In this study the spatial and temporal relationships, i.e. teleconnections between the oceanic climate states and the regional weather in Thailand are assessed by various techniques of stochastic time series analysis. Time series of the sea surface temperature (SST) and various ocean indices of the Pacific and the Indian Oceans as well as of climate variables of 121 meteorological stations from 5 regions across Thailand that include humidity, evaporation, temperature and rainfall during 1971-2007 are examined using cross-correlation, linear-regression and wavelet transform methods. The analysis concentrates on meteorological observations in the eastern seaboard of Thailand which has particularly been suffering from water shortage problems in recent years.

The results of the statistical analyses show that the El-Niño 1+2 SST index of the eastern Pacific Ocean correlates the strongest with the Thai local climate. The most sensitive parameters to the ocean indices are the minimum temperature at stations in the northern and northeastern, inland regions of Thailand, and the number of rainy days in the eastern, central and southern, coastal regions of Thailand. In the south the rainfall varies positively with El-Niño along the Gulf of Thailand, but negatively along the Andaman Sea, with maximal correlations at lag-times of 2-4 months. The strongest influence of El-Niño 1+2 on the local climate is in the pre-monsoon season (Jan-Mar).

Using the results of the relationships between ocean indices and local climate as well as of downscaled predictions of GCM- models, multiple linear dynamic regression models as well as ARMA- and ARMAX- (with exogenous variables) models are set up to forecast the local climate variables for the Eastern Seabord (EE) region of Thailand in the medium- and long-term. For the minimum temperatures, in particular, for both the regression- or an ARMA-model is used, the forecast reliability is improved significantly when the El Niño 1+2 SST’s are used as an additional external regressor or exogenous variable in the model.
1. INTRODUCTION
Most of Thailand’s the water supply for urban, agricultural and industrial use is surface water stored in scattered large-scale regional reservoirs. Considering that much of Thailand is a peninsula located between the Pacific and Indian Ocean (Fig. 1), it is no wonder that most of its regional weather as well as its water resources are affected by the variations of the ocean-driven monsoons. The long coastline of Thailand which in the eastern and western part is connected to the Gulf of Thailand and the Andaman Sea, respectively, conveys the wet and warm air northward which, in turn, have pretty much dried up when reaching the northern, mountainous regions of Thailand. Because of this peculiar geographic location, Thailand’s water resources are strongly influenced by the mix of tropical wet, tropical dry and tropical monsoon seasons. The climate patterns over the various regional watersheds, where much of the rainfall is accumulated in reservoirs, fluctuate significantly, depending on the movement of the monsoon. Thus these reservoirs usually gather about 80% of the annual rainfall during the 6 months of the rainy season (May-October) when the monsoons are moving northward. During the dry season (November - April) the winds are blowing from the northern Indo-China land mass and rain drops only a few days in a month.

Since the monsoons, in turn, depend themselves upon the thermal states of the Pacific (Walker, 1924; Singhrattna et al., 2005a) and the Indian Ocean (Clark et al., 2000), the local climate not only in Thailand, but also in southeast Asia as a whole, is indirectly affected by these ocean states. These teleconnections act not only on the seasonal but on the inter-annual scale as well (Bejranonda and Koch, 2010c), as is witnessed by the extreme weather change pattern with long periods of heavy droughts Thailand has seen in recent years (Bejranonda et al., 2010a). In the present paper the possible relationships between the ocean states and the local weather (e.g. Walker, 1924; Pant and Parthasarathy, 1981) will be investigated by analyzing, on one hand, various ocean state parameters, namely, the and sea surface temperatures (SST) and, on the other hand, historical climate data (precipitation and temperatures) across Thailand, to understand and quantify the connections between these two groups.

There have been numerous studies up-to-date of the impacts of ocean-state indicators, such as ENSO and SOI, on the earth’s climate system, in general (e.g. Glantz, 2001), as well as on its wide impacts on regional hydrological systems (e.g. Kiladis and Sinha 1991; Chen and Kumar, 2002; Samuel et al., 2006). The influence of the oceans on the monsoons with subsequent seasonal and annual variations of precipitation and temperatures in tropical regions has been clearly demonstrated by (Ramage, 1971), Krishnamurti et al. (1989a, 1989b) and Webster et al. (1998). Relationships between the local climate in southeast Asia and the ocean states with the Indian monsoon as a mediator have been established by Rasmusson and Carpenter (1983), Fein and Stephens (1987), Webster et al. (1998), Krishna Kumar et al. (1999) and Krishnamurthy and Goswami (2000) as well as for the China monsoon by Yu et al. (2001), Chan and Zhou (2005) and Wang et al. (2008).

Although the monsoon variability over central Thailand has been studied by Singhrattna et al. (2005a), there are still other various adjacent indices over both the Pacific and Indian oceans that appear to relate to variations of the Asian monsoon (Feng, 2001). However, the exact temporal relationships, as well as their spatial variations across Thailand are still waiting to be examined. This is the major objective of the present paper. More specifically, the spatial and temporal relationships between the oceanic climate states and the regional
Oscillation Index (SOI), (6) the Extreme Eastern Tropical Pacific SST (Niño 1+2), (7) the detrimental effects on the availability of hydrological water resources in Thailand. A pilot ARMAX- models for the purpose of medium- and long-term climate prediction and forecasting of the local climate in the region.

The climate variables are 1971-2007 monthly time series of average evaporation, average temperature, average daily rainfall, number of rainy days and monthly rainfall from 121 meteorological stations, administered by the Thai Meteorological Department, and scattered across 5 regions of Thailand: North (NN), Northeast (NE), East (EE), Central

2. TIME-SERIES ANALYSIS

2.1 Ocean indices and climate data
The monthly atmospheric and ocean time series of the sea surface temperature (SST) and the various other ocean state indices, whose denominations and locations are shown in Fig. 1, are taken from the Ocean Observations Panel for Climate (OOPC, 2009) which provides calculated indices using the method of Reynolds OIv2 SST analysis (Reynolds, 1988; Reynolds and Marsico, 1993). The three Indian Ocean SST’s indices are (1) the Southeastern Tropical Indian Ocean SST index (SETIO), (2) the South Western Indian Ocean SST index (SWIO) and (3) the Western Tropical Indian Ocean SST index (WTIO). As for the Pacific Ocean indices, provided by NOAA (CPC, 2009), they are (1) the Eastern Pacific Oscillation (EP), (2) the Northern Oscillation Index (NOI), (3) the Pacific Decadal Oscillation (PDO), (4) the Pacific North American Index (PNA), (5) the Southern Oscillation Index (SOI), (6) the Extreme Eastern Tropical Pacific SST (Niño 1+2), (7) the Eastern Tropical Pacific SST (Niño 3), (8) the Central Tropical Pacific SST (Niño 4) and (9) the East Central Tropical Pacific SST (Niño 3.4).

The climate variables are 1971-2007 monthly time series of average evaporation, average relative humidity, extreme maximum temperature, extreme minimum temperature, maximum rain in 24hr, mean maximum temperature, mean maximum temperature, mean temperature, average daily rainfall, number of rainy days and monthly rainfall from 121 meteorological stations, administered by the Thai Meteorological Department, and scattered across 5 regions of Thailand: North (NN), Northeast (NE), East (EE), Central

The 1st Symposium on Sustainable Water Management and Climate Change Adaptation co-organized by Nakhon Pathom Rajabhat University and The University of Kassel, Germany
Nakhon Pathom Rajabhat University, Nakhon Pathom, Thailand, June 16 - 17 , 2010.
(CC) and South (SS) (Fig. 1). In addition to the observed climate data, predictors of the “Third Generation Coupled Global Climate Model” (CGCM3) (CCCMA, 2009) from the Canadian Centre for Climate Modeling and Analysis (CCCMA) are selected to improve the medium- and long-term climate prediction across the pilot region in the eastern seabord of Thailand.

2.2 Wavelet analysis
The time-series above are examined using the continuous wavelet transform (CTW) to look for temporal changes of their power spectrum. As described in Torrence and Compo (1998), the wavelet transform can be somewhat considered as a time-slided windowed Fourier transform (FT), though with advantageous resolution properties in the time- and scale (frequency)- domain. Thus, unlike in the regular FT, in the CWT high frequencies (low scales) in the signal are well resolved in time, but poorly in Fourier space, and with the opposite holding for low frequencies (Kumar and Foufoula-Georgiou, 1997). The CTW has been used frequently in recent times for the long-term analysis of hydro-meteorological and climate data (e.g. Gaucherel, 2002; Markovic and Koch 2005, Koch and Markovic, 2007). Depending on the structure of the time series, different wavelet functions can be employed. Here the relatively smooth Morlet wavelet is used.

The local (2D) wavelet power spectra (scalograms) for three climate variables from the EE- pilot region as well as for three Niño SSTs are shown in Fig. 2. One can notice from these scalograms that the ocean states visually correlate with the observed climate variables in this region of Thailand. This holds especially for the strong Niño 1+2 SST peak (located in the eastern Pacific, see Fig. 1) in the 1995-2000 time period. As discussed by Markovic and Koch (2005), these connections could be further quantified by the calculation of the wavelet cross-correlation spectrum in the frequency domain. However, as discussed in the following section, a more intuitive way is the use of the classical cross-correlation in the time-domain.

2.3 Cross-correlation
Serial correlation, cross-correlation or spatial correlation (Yue, 2002) of two time series $X$ and $Y$ from different hydrological sites can be used to infer teleconnective relationships between the two. In particular, by lagging one of the time series by time $\tau$ during the cross-correlation, the optimal delay time $\tau_{\text{opt}}$ for which the cross-correlation function $R(\tau)$,

$$R(\tau) = \frac{E[(X_{t+\tau} - \mu_X)(Y_t - \mu_Y)]}{\sigma_X \sigma_Y}$$

(1)

(where $\mu$ are the mean, and $\sigma$ the standard deviations of the corresponding time series) is maximum, i.e. $R_{\text{max}} = R(\tau_{\text{opt}})$, can be found.

The results of the application of Eq. (1) to the correlations of the vector of the four Niño SSTs (located in Fig. 1) with the vector of the three climate variables at the meteorological station 478201 in the EE-region is shown in Fig. 3. One notes that for this pilot station, surprisingly, the maximum correlations $R_{\text{max}} = R(\tau_{\text{opt}})$ are obtained with the Niño 1+2 SST, located the farthest away in the eastern Pacific, at a lag of $\tau_{\text{opt}} = -3$ months, whereas the other two ocean indices, Niño 3 and Niño 3.4, located closer to Thailand, have lags of $\tau_{\text{opt}} = -2$ and -1 months, respectively, though with smaller values of $R_{\text{max}}$. 

The 1st Symposium on Sustainable Water Management and Climate Change Adaptation co-organized by Nakhon Pathom Rajabhat University and The University of Kassel, Germany Nakhon Pathom Rajabhat University, Nakhon Pathom, Thailand, June 16 - 17 , 2010.
Figure 2: Wavelet power spectra of measured local climate variables in the EE region of Thailand and of the Niño Sea surface temperatures during 1974-2006.

Figure 3: Cross-correlations of the four Niño SSTs with local monthly minimum, maximum temperature and rainfall at station 478201 in the EE-region during 1971-2006.
The maximum correlations $R_{\text{max}} = R(\tau_{\text{opt}})$ for all 13 climate stations located in the EE-region with the 12 different Indian and Pacific Ocean state variables (defined in the previous section) are listed in Table 1. One can observe that the Niño SSTs, namely, Niño 1+2 SST correlate best with the local climate variables in that region of Thailand.

Table 1: Average cross-correlation coefficients between ocean indices and the basic meteorological time series of 13 stations in the EE-region. The highest positive and negative correlation coefficients are accentuated with green and red colors, respectively.

<table>
<thead>
<tr>
<th>Indices</th>
<th>avg.evap</th>
<th>avg-rh</th>
<th>extr.max</th>
<th>extr.min</th>
<th>max.rain</th>
<th>24hr</th>
<th>mean.max</th>
<th>mean.min</th>
<th>mean.temp</th>
<th>mean.temp</th>
<th>Rain mm/day</th>
<th>rainy-day</th>
<th>total.rain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indian-SETIO</td>
<td>0.02</td>
<td>0.00</td>
<td>0.14</td>
<td>0.09</td>
<td>-0.02</td>
<td>0.17</td>
<td>0.09</td>
<td>0.13</td>
<td>0.13</td>
<td>0.00</td>
<td>0.05</td>
<td>-0.01</td>
<td></td>
</tr>
<tr>
<td>Indian-SWIO</td>
<td>-0.14</td>
<td>-0.02</td>
<td>0.07</td>
<td>0.04</td>
<td>0.03</td>
<td>0.08</td>
<td>0.06</td>
<td>0.05</td>
<td>0.00</td>
<td>0.07</td>
<td>0.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indian-WTIO</td>
<td>-0.07</td>
<td>0.00</td>
<td>0.20</td>
<td>0.13</td>
<td>-0.02</td>
<td>0.23</td>
<td>0.14</td>
<td>0.19</td>
<td>-0.04</td>
<td>-0.02</td>
<td>-0.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pacific-EP</td>
<td>-0.20</td>
<td>0.44</td>
<td>-0.20</td>
<td>0.47</td>
<td>-0.07</td>
<td>-0.43</td>
<td>0.50</td>
<td>0.35</td>
<td>0.22</td>
<td>-0.09</td>
<td>-0.42</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pacific-Nino1.2</td>
<td>0.34</td>
<td>-0.61</td>
<td>0.43</td>
<td>-0.15</td>
<td>-0.34</td>
<td>0.64</td>
<td>0.24</td>
<td>0.58</td>
<td>-0.21</td>
<td>-0.76</td>
<td>-0.42</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pacific-Nino3</td>
<td>0.35</td>
<td>-0.40</td>
<td>0.45</td>
<td>-0.19</td>
<td>-0.30</td>
<td>0.60</td>
<td>-0.04</td>
<td>0.60</td>
<td>-0.17</td>
<td>-0.63</td>
<td>-0.36</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pacific-Nino3.4</td>
<td>0.25</td>
<td>-0.35</td>
<td>0.29</td>
<td>0.13</td>
<td>-0.22</td>
<td>0.46</td>
<td>0.25</td>
<td>0.48</td>
<td>-0.13</td>
<td>-0.42</td>
<td>-0.31</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pacific-Nino4</td>
<td>0.13</td>
<td>-0.16</td>
<td>0.20</td>
<td>0.32</td>
<td>-0.15</td>
<td>0.23</td>
<td>0.31</td>
<td>0.33</td>
<td>-0.03</td>
<td>-0.31</td>
<td>-0.25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pacific-NOI</td>
<td>-0.04</td>
<td>0.00</td>
<td>-0.14</td>
<td>-0.04</td>
<td>0.06</td>
<td>-0.16</td>
<td>-0.09</td>
<td>-0.18</td>
<td>0.01</td>
<td>0.10</td>
<td>-0.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pacific-PDO</td>
<td>0.12</td>
<td>-0.06</td>
<td>0.18</td>
<td>0.18</td>
<td>0.02</td>
<td>0.20</td>
<td>0.17</td>
<td>0.20</td>
<td>0.05</td>
<td>-0.16</td>
<td>-0.06</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pacific-PNA</td>
<td>-0.02</td>
<td>-0.06</td>
<td>0.06</td>
<td>0.06</td>
<td>-0.06</td>
<td>0.08</td>
<td>0.12</td>
<td>0.17</td>
<td>-0.07</td>
<td>-0.01</td>
<td>0.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pacific-SOI</td>
<td>-0.09</td>
<td>0.08</td>
<td>-0.12</td>
<td>-0.11</td>
<td>0.09</td>
<td>-0.20</td>
<td>-0.13</td>
<td>-0.19</td>
<td>0.01</td>
<td>0.13</td>
<td>0.10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pacific-WP</td>
<td>0.06</td>
<td>-0.13</td>
<td>0.07</td>
<td>0.02</td>
<td>-0.09</td>
<td>0.03</td>
<td>0.06</td>
<td>0.01</td>
<td>-0.09</td>
<td>-0.05</td>
<td>-0.07</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 4: Temporal and spatial correlations between Niño 1+2 SST and mean minimum temperatures across Thailand. Left panel: maximum correlation coefficients; right panel: lag-times at correlation maximum.
While Table 1 lists only the average maximum cross-correlation coefficients $R_{\text{max}}$ of the EE-pilot region, the maps of Fig. 4 illustrate the variations of $R_{\text{max}}$ (left panel), together with the corresponding optimal lag-time $\tau_{\text{max}}$ (right panel) for the correlations of the mean minimum temperature of all climate stations across Thailand with the Niño 1+2 SST which turned out to have the highest correlation with the local climate across the country. From the left map of Fig. 4 one notices the strong spatial teleconnection of this Pacific ocean state variable, whereby stations in the east (EE) and south of the country exhibit positive correlations, while those in the center and the north have negative correlations. The temporal teleconnections between Niño 1+2 SST and minimum temperatures are indicated in the right panel of Fig. 4 by the optimal lag time $\tau_{\text{opt}}$ with the highest correlation $R_{\text{max}}$. One notes a peculiarity for most of the stations in the eastern seabord region EE. Unlike for the stations in the center, east and north of Thailand, which show lag times of +2 to +4 months, those in the EE region have optimal lag times of -2 to -4 months, similar to most of the stations on the Pacific-coastline of south Thailand.

2.4 Time-lagged regression analysis
Another option to visualize and quantify the teleconnective relationships between the time series of the ocean state variables $x_t$ and of the climate variables $y_t$ consists in the application of time-lagged linear regression analysis in the form

$$y_t = x_{t+\tau} * \beta_1 + \beta_0 + \epsilon,$$  \hspace{1cm} (2)

where $x_{t+\tau}$ is the time-lagged (by $\tau$ months) series of the ocean state predictor variable $x_t$. The regression coefficients $\beta_1$ and $\beta_0$ are found by minimization of the error term $\epsilon$ using the method of least-squares. The goodness of the regression fit is then measured by the coefficient of determination $R^2$ defined as

$$R^2 = 1 - \frac{\sum_{t=1}^{n} (y_t - \bar{y})^2}{\sum_{t=1}^{n} (y_t - \bar{y}_T)^2},$$  \hspace{1cm} (3)

where $\bar{y}$ is the mean of $y$ and $\bar{y}_T$ is the predicted value.

Eq. (2) is solved repeatedly for different lag-times $\tau$ of the observed ocean variable to find the optimal lag-time $\tau_{\text{opt}}$ which best predicts the climate variable at a particular station. This has been done not only for the total of the measured monthly values (annual series), but for the split data sets of the four different seasons of the year, i.e. the dry season [Oct-Dec], the pre-monsoon season [Jan-Mar], the monsoon1 season [Apr-Jun], the monsoon2 season [Jul-Sep], as well. Using this seasonal separation, the important influence of the summer monsoon on the water resources in Thailand can be better worked out (Singhrattna et al., 2005b).

The results of this analysis for the observed minimum temperature at the EE station 478201 regressed on the lagged El Niño 1+2 SST data are shown in Fig. 5. One notes that, depending on the lag-time $\tau$ used, the predictor/predictand point pairs $(x_{t+\tau}, y_t)$ are more or less scattered around the linear regression line. Corroborating the results of the correlation analysis of the previous section, the linear relationship between the El Niño 1+2 SST and the station minimum temperature is satisfied best (as quantified by the largest value of $R^2$) for the optimal lag-time $\tau_{\text{opt}} = 3$ months.
As for the important seasonal relationships between the SST and the minimum temperatures, for the pre-monsoon1 season, Jan-Mar, for example, the optimal lag time is at $\tau_{opt} = -1$ months. Similar results are obtained for the regressions of the maximum temperatures and of the various variables used to quantify the station rainfall (see Table 1).

3. MEDIUM- AND LONG-TERM CLIMATE PREDICTION USING TELE-CONNECTIONS AND CLIMATE MODELS
Because of the precarious water-availability situation in Thailand which, in the long run, will additionally be most likely exacerbated by imminent climate change in that region of the world (IPCC, 2007; Koch, 2008), the previously established ocean-Thailand seasonal-
weather-pattern teleconnections can provide a tool for possible forecasting of short-term seasonal climate pattern across the country (Webster et al., 1998) and so to better manage the adequate supply of surface water resources to the users (Bejranonda and Koch, 2010a). To extend the forecasting to a longer time-scale, which is of particular interest for water planners, GCM-predictors from global climate models must be used. This is the objective of the present section.

3.1 Multiple linear regression model with GCM- and dynamic ocean predictors
To predict the monthly climate variables at a meteorological station, a multiple linear regression model, called a transfer model in the language of statistical downscaling - for the climate variable in terms of the GCM-predictors is set up. To improve the prediction power of this downscaled GCM-model - details of which are discussed in Bejranonda and Koch (2010a), namely the choice of the best predictors - it has been enlarged by the extra inclusion of the optimally time-lagged ocean state predictors, in the case of which the multiple linear regression equations may be considered as a dynamic regression model that can be written as

\[ y = x_1\beta_1 + \cdots + x_p\beta_p + [I'_{\text{index}}] + c \]  

Here \( y \) is the dependent climate variable vector; \( x \), the independent GCM predictor vector; \( \beta \), the vector of the GCM- predictors and the term \( I'_{\text{index}} \) denotes the regression coefficients of the optimally lagged ocean index \( I \). The latter are shut off and on in the subsequent analysis to understand their benefits in the prediction of the climate variables.

The prediction-performance of the various model variants is measured by the Nash–Sutcliffe (NS) model efficiency \( E \) (Nash and Sutcliffe, 1970) which is determined during the verification process and which is defined as

\[ E = 1 - \frac{\sum_{t=1}^{T}(Q^o_t - Q^m_t)^2}{\sum_{t=1}^{T}(Q^o_t - \bar{Q^o})^2} \]  

where \( Q^o_t \) is the observed value at time \( t \), \( Q^m_t \) is modeled value at time \( t \) and \( \bar{Q^o} \) is the mean of observed values.

The coefficients of the regression model above are determined by least squares, employing the time series data of the meteorological variables and the CGCM3 predictors of the macro-scale regional climate variations for the EE- study region for the 1961-2000 calibration period. In case that the regression model is augmented by the ocean state predictors, the observed ocean data are incorporated as well. Using these calibrated predictors, the temperatures and rainfall at selected stations are then forecasted for the verification period 2001- 2007.

3.2 ARMA and ARMAX models
3.2.1 Autoregressive moving average (ARMA) models
Another option for forecasting stochastic time series consists in the use of the so-called Box-Jenkins approach (Box and Jenkins, 1970) which leads to the concept of the autoregressive - moving average (ARMA) models. ARMA models are powerful tool in economic forecasting and are increasingly being employed also in hydrological and geophysical time series prediction (e.g. Sun and Koch, 1996). Unlike the dynamic
regression models used in the previous section, the ARMA model is build from the structure of the time series itself, so that, eventually, the future behavior of the time series is predicted from past values alone.

The most general ARMA\((p,q)\) - model for a mean-adjusted time series \(Y_t\) \((t=1,\ldots,n)\) i.e. \(y_t = Y_t - \bar{Y}\) with a \(p\)-order autoregressive (AR) and a \(q\)-order moving average (MA) process can be written as

\[
y_t = \sum_{i=1}^{p} \phi_i y_{t-i} + \sum_{i=1}^{q} \theta_i \epsilon_{t-i} + \epsilon_t, \tag{6}
\]

where \(\phi_i (i=1,\ldots,p)\), \(\theta_i (i=1,\ldots,q)\), are the unknown coefficients of the AR and MA-process, respectively, and \(\epsilon_t\) is the noise which is assumed to be uncorrelated.

A generalization of an ARMA- model is the ARIMA-model which is applicable when the time-series is non-stationary. Then an initial differencing step of order \(d\) which depends on the kind of non-stationarity and which corresponds to the integrated (I) part of the model, is required to remove the non-stationarity. Such a model is generally referred to as an ARIMA\((p,d,q)\) - or ARMA\((p,d,q)\)- model. A further extension is obtained when the time-series under question exhibits seasonality with a certain period, say \(S\). This is often the case for hydrological time series so that for a monthly series the seasonal period would be \(S=12\), i.e. one year. In such a case the ARIMA\((p,d,q)\)- model is replaced by a so-called multiplicative SARIMA\((p, d, q)\times(P,D,Q)_S\) - model, where the triplet \((P,D,Q)_S\) denotes the orders of the seasonal AR, of the seasonal differencing, and of the seasonal MA, respectively (e.g. Box and Jenkins, 1970; Brockwell and Davis, 1996).

As prescribed by the Box-Jenkins formalism, the build-up of the general ARMA- or ARIMA- model for a time series \(y_t\) requires essentially three stages that are (1) identification, (2) estimation, and (3) diagnostic checking. The first stage involves the computation of the total and partial autocorrelations; a graphical check of the latter to determine the positions of spikes and their general decay pattern for various lag-times \(y_{t-k}\) \((k=1,\ldots,p)\); and the application of various \(t\)-tests of statistical significance to ascertain which lag-coefficients do not contribute anymore to the tentative ARMA model. In the second step the coefficients \(\phi_i\) and \(\theta_i\), of the ARMA-model are then estimated by nonlinear minimization of the maximum-likelihood function for the prediction error. For the special case of a pure AR-model (no \(\epsilon\)-term in Eq. (6)) the model estimation reduces to a linear least squares problem for the so-called Yule-Walker equations. In the third step the quality of the estimation and the validity of the assumed model is checked \textit{a posteriori} by evaluating the AIC (Akaike information criterion) or the BIC (Schwarz information criterion) and by post-inspecting the various autocorrelation functions for the model residuals (e.g. Box and Jenkins, 1970; Brockwell and Davis, 1996).

Once the structure of the ARIMA-model has been determined, so-called \textit{“k- step ahead”} – forecasts \(\hat{y}_{n+k}\) of the time-series after the end of the observation period \(t_n\) can be carried out, whereby Eq. (6) is solved recursively for time \(t_{n+k}\) by plugging in either the known (observed) \(y_{n-i}\) \((i=1,\ldots,n-1)\) or – once the observation horizon is passed – the already predicted values \(\hat{y}_{n+k-j}\) \((j=1,\ldots,l-1)\) from the previous forecast time step \(t_{n+k-1}\). From the theoretical forecast formulation error estimates (variances \(\text{Var}(\hat{y}_{n+k})\)) and, eventually,
confidence intervals for the forecasts $\hat{y}_{n+k}$ can be determined. It is clear that $\text{Var}(\hat{y}_{n+k})$ increases largely with increased $k$ of the ahead-prediction horizon $n+k$. In fact, theory shows that the upper and lower confidence half-widths scale roughly with the square root of the forecast horizon $k$. This behavior leads to the typical shape of the confidence interval, as shown later in Fig. 6.

3.2.2 ARMA models with exogenous variables (ARMAX-model)
A further extension and improvement of the classical ARMA-model is the incorporation of an external predictor (exogenous variable) $x$ in the forecast model for the climate variable $y$. This leads to the so-called ARMAX-model (Hipel and McLeod, 1993) which is obtained by simply extending Eq.(6) to

$$y_t = \sum_{i=0}^{q} \phi_i y_{t-i} + \sum_{i=0}^{r} \varphi_i x_{t-i} + \sum_{i=1}^{q} \theta_i \epsilon_{t-i} + \epsilon_t, \quad (7)$$

where $\varphi_i$ ($i=0,\ldots,r$), are the new coefficients of the exogenous predictor process $x$. Note that the index for $x$ starts here at $i=0$ to allow for the possibility of instantaneous (no-lag) response of $y$ to $x$. In the present application the exogenous variables $x$ are either the GCM-predictors or the ocean state predictors which, as previously shown, correlate reasonably well with the observed climate variables across Thailand. Other than that follows the establishment of the ARMAX-model the same three-step procedure as that of the ARMA-model.

3.3 Results
Fig. 6 and Table 2 illustrate some results obtained with the various time-series models, namely, (1) dynamic multiple regression using GCM-predictors, (2) dynamic multiple regression with GCM-predictors plus external ocean SST regressors, (3) the pure ARMA-(ARIMA) model, (4) the ARMAX-model using the ocean SST as exogenous variable, and (5) the ARMAX-model with exogenous GCM-predictors.

Fig. 6 shows the minimum temperature forecasts at the EE-pilot station for the 2001-2006 verification period, using the pure ARMA-model and the two ARMAX-model variants. One clearly notes a gradual improvement in the forecast, measured by the NS-coefficient $E$ (Eq. 5), as $E$ increases from $E=0.773$ for the pure ARMA-model, to 0.82 for the ARMAX-GCM, to 0.84 for the ARMAX-SST model.

Table 2 lists the NS-coefficients $E$ of the prediction performances for all three climate variables at the EE-pilot station for the two variants of the dynamic regression models. One notes that, for the minimum temperature at least, the ARMA-models are also better than the multiple linear regression models. On the other hand, for the maximum temperature but, particularly for the average monthly rainfall, the advantages of using the incrementally-becoming-more-complex time series models are not obvious. Eventually, this is due to the fact that the rainfall in the region can neither be well predicted by an GCM-model, nor does it correlate well with the various ocean state indices (see Table 1).
Figure 6: Verification of the minimum temperature prediction at the EE-pilot for the 2001-2006 period using the pure ARMA(p, d, q) \times (P,D,Q)_S model and the two ARMAX-models (see text). The orange and the yellow bands denote the 90% and the 95% confidence intervals for the prediction, respectively.

Table 2: Model performance, as measured by the Nash-Sutcliffe efficiency coefficient $E$, for min/max temperature and rainfall in the EE region for the verification period 2001-2006 using the various downscaled CGM- (CGCM3-T47 and CGCM3-T47 model) multiple linear regression models (Eq. 4), without and with El Niño 1+2 SST predictors included, as well as of ARMA- (Eq.6) and the ARMAX-SST (Eq. 7) models.

<table>
<thead>
<tr>
<th>Models and data</th>
<th>Nash-Sutcliffe model efficiency $E$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>min temp</td>
</tr>
<tr>
<td>linear regression with CGCM3-T47</td>
<td>0.67</td>
</tr>
<tr>
<td>linear regression with CGCM3-T47 and Niño 1+2 SST</td>
<td>0.68</td>
</tr>
<tr>
<td>linear regression with CGCM3-T63</td>
<td>0.51</td>
</tr>
<tr>
<td>linear regression with CGCM3-T63 and Niño 1+2 SST</td>
<td>0.60</td>
</tr>
<tr>
<td>ARMA- model</td>
<td>0.77</td>
</tr>
<tr>
<td>ARMA- model with Niño 1+2 SST exogenous variable</td>
<td>0.84</td>
</tr>
</tbody>
</table>

4. CONCLUSIONS
Spatial and temporal relationships, i.e. teleconnections between the oceanic climate states and the regional weather in Thailand are assessed by various techniques of stochastic time series analysis. The results show that the El-Niño 1+2 SST anomaly index of the Pacific Ocean, which refers to the most eastern section of the Pacific, correlates the strongest with Thailand’s local climate. The classical cross-correlation shows that the most sensitive parameters to the ocean indices are the minimum temperature at stations in the northern and northeastern, inland regions of Thailand, and the number of rainy days in the eastern, central and southern, coastal regions of Thailand. In the southern region the amount of
rainfall at the coast of Gulf of Thailand varies positively with El-Niño, but negatively for stations along the Andaman Sea coast in the west of the isthmus, with maximal correlation lag-times of 2-4 months. The strongest influence of El-Niño 1+2 plays the role over local climate on the pre-monsoon season, during Jan-Mar, in Thailand.

Using the results of the relationships between ocean indices and local climate as well as of downscaled predictions of GCM- models, multiple linear dynamic regression models as well as ARMA and ARMAX- models (with external exogenous variable) are set up for the medium- and long-term forecast of some of the local climate variables across Thailand. The results show that, at least for the minimum temperatures in the EE pilot region, regardless of whether the regression- model or an ARMA-model is used, the forecast reliability is improved significantly when the El Niño 1+2 SST’s are used as an additional external regressors in the regression model or as an exogenous variable in the ARMA-model (=ARMAX model). As is typical for the forecast of rainfall in climate studies, this important climate variable is not well predicted in our study as well by the methods used.

In any case, the results of our analysis indicate the possibility of a better forecast of extreme seasonal climate variations across some regions of Thailand over a limited time period by using expected medium-term variations of the Pacific and Indian ocean indices. Furthers studies are needed. Though, to validate the forecasts for other regions of Thailand than the ones investigated here

ACKNOWLEDGMENTS
The authors wish to thank Dr.Sucharit Koontanakulvong at Department of Water Resources Engineering, Chulalongkorn University for assisting with the data compilation.

REFERENCES


Singhrattna N., B. Rajagopalan, K. Krishna Kumar and M. Clark (2005a) Interannual and interdecadal variability of Thailand summer monsoon, J. Climate, 18, 1697-1708.


Walker, GT. (1924) Correlation in seasonal variations of weather, IV, a further study of world weather, Memories of the Indian Meteorological Department, 24, 275–332.


Yu, R.C., M.H. Zhang, Y.Q. Yu and Y. M. Liu (2001) Summer monsoon rainfalls over Mid-Eastern China lagged correlated with global SSTs, Advances in Atmospheric Sciences, 18, 179-196