Measuring Recovery to Build up Metrics of Flood Resilience Based on Pollutant Discharge Data: A Case Study in East China

Jinglu Song 1, Bo Huang 1,2,3,* and Rongrong Li 2

1 Department of Geography and Resource Management, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong; jinglusong@link.cuhk.edu.hk
2 Institute of Space and Earth Information Science, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong; rongrongli@cuhk.edu.hk
3 Shenzhen Research Institute, The Chinese University of Hong Kong, Shenzhen, China
* Correspondence: bohuang@cuhk.edu.hk

Received: 14 June 2017; Accepted: 15 August 2017; Published: 18 August 2017

Abstract: Building “disaster-resilient” rather than “disaster-resistant” cities/communities requires the development of response capabilities to natural disasters and subsequent recovery. This study devises a new method to measure resilience via recovery capability to validate indicators from social, economic, infrastructural, and environmental domains. The pollutant discharge data (waste-water and waste-gas discharge/emission data) of local power plants, sewage treatment plants and main factories were used to monitor recovery process of both people’s living and local industrial production as the waste water/gas is released irregularly during the short disaster-hit period. A time series analysis of such data was employed to detect the disturbance on these infrastructures from disasters and to assess community recovery capability. A recent record-breaking flash flood in Changzhou, a city in eastern-central China, was selected as a case study. We used ordinal logistic regression to identify leading proxies of flood resilience. A combination of six variables related to socioeconomic factors, infrastructure development and the environment, stood out and explained 61.4% of the variance in measured recovery capability. These findings substantiate the possibility of using recovery measurement based on pollutant discharge to validate resilience metrics, and contribute more solid evidences for policy-makers and urban planners to make corresponding measures for resilience enhancement.

Keywords: recovery capability; disaster resilience; resilience measurement; pollutant discharge data; East China

1. Introduction

The increased frequency of record-breaking flash flood hazards has become a widespread problem in Chinese cities [1]. In addition to building damage, urban flooding depresses property prices, creates health risks, degrades water quality, and has an ongoing effect on local production and daily life [2]. Cities prone to flood risks have taken measures to improve disaster resistance and rescue work [3]. However, due to the global climate change, the increased frequency of flood hazard followed by a high-intensity rainfall event renders conventional measures with emphasis on flood resistance and rescue work inadequate, and makes a comprehensive study necessary, particularly on post-disaster recovery. Recent notions of building “disaster-resilient” rather than “disaster-resistant” cities/communities in developed countries [4–6] could complement traditional disaster management approaches. In addition to the preparedness of communities involved in the context of disaster resistance, the concept of disaster resilience emphasizes the response of communities during a disaster and their recovery measures afterwards especially for extreme events communities that
cannot reasonably resist. Compared to disaster resistance measures, the resilience measures developed for a specific to one type of disasters could benefit other threats and thus reduce the total cost that may be incurred. Table A1 (in Appendix A) gives the detailed comparisons between disaster resistance and disaster resilience.

Numerous studies have addressed how to build or enhance disaster resilient communities efficiently, and the capability of measuring resilience has been recognized as the next necessary step [7,8]. Based on quantitative measurement of resilience, policy makers could understand dominant factors that make some communities more resilient than others, or being more resilient over time, and could help them explore the most effective risk reduction efforts that could move the community toward disaster resilience. Unfortunately, as there is no agreed definition of disaster resilience, directly measuring resilience is challenging [9,10]. Most attempts to quantitatively measure disaster resilience define a set of characteristics or attributes previously identified through practical experience or research [7,11–15]. The characteristics of resilience usually include a set of attributes representing various dimensions, such as social, economic, infrastructural, and so on.

While the obvious advantage of this approach is its adaptability to distinct circumstances, geographical settings, or cultures, the selection of variables representing different characteristics as a proxy of resilience measurement remains subjective and the contributions of these variables in developing resilience has to be validated [9]. Using external measurement (e.g., post-disaster population increase, indicator of building reconstruction, etc.) represents a useful approach to validate the choice of variables [8,16]. The term recovery is widely used in the conceptualizations of disaster resilience [17]. Most scholars define the concept of resilience in terms of the speed or how quickly a community or a society can recover from a disaster [18–22]. This is because evidence has shown that recovery processes are closely related to antecedent conditions which serve as a baseline for resilience assessment. The antecedent conditions are determined by multi-dimensional community characteristics, such as social, economic, infrastructural, environmental components, each of which is associated with the metrics of resilience measurement [11,23,24]. The higher the resilience, the more quickly a community recovers to normal functioning [25,26]. It is within the context that recovery outcomes can serves as a resource and practical lessons to inform communities of the key intrinsic factors that contribute to disaster resilience [27]. Therefore, some measures tend to focus on response and recovery outcome after the damaging events for the validation of resilience measurement [8,16,28]. Based on these empirical works, useful recovery indicators and models have proved to be important decision support tools for increasing disaster resilience and reducing disaster vulnerability.

Since defining the concept of recovery should specify the type and intensity of disaster (e.g., flood or earthquake, catastrophic disaster or minor event) [17,24,29] and distinguish short-term recovery from long-term recovery [30–32], selecting proper indicators as a proxy of community recovery performance should be particularized case by case. For disasters with short-term recovery process, post-disaster community performance mainly focus on restoration to pre-disaster functions, and at most cases are evaluated by a set of infrastructure recovery indicators, such as building reconstruction, refunctions of key utilities and productivity recovery [33,34]. While the indicator used to measure building reconstruction has been widely documented as a proxy of community recovery performance in a number of literatures [8,35], few empirical studies have been conducted to build recovery indicators of key utilities and regional productivity, though scholars concern reffunctioning of people’s lives and livelihoods more than the reconstruction of buildings [30,36,37]. This is mainly due to the data unavailability for measuring such recovery indicators, especially in developing countries [38], which incurs an ongoing challenge in building up external-validated metrics of flood resilience [39,40].

To narrow this gap, this study proposes a new method to measure short-term recovery capability, using the recovery measurement as external validation term of flood resilience metrics. We select as a case study a recent flash flood in Changzhou, a city in eastern-central China. We attempt to address the following problems: (1) measuring recovery capability after flood hazards within a short term based on time analysis of waste-water and waste-gas discharge and emission
data; and (2) finding a set of variables that are validated by recovery measurement and can provide the best assessment of multi-dimensional characteristics of flood resilience. The paper is organized as follows. Section 2 provides a brief description of the study area and details the procedures of measuring recovery capability. Section 3 describes how to link the measured recovery capability to resilience. Sections 4 and 5 examine the results of the recovery measurement and their influence on the choice of variables for measuring resilience. The final section concludes with a discussion of the main findings and limitations.

2. Methods

2.1. Study Area

From 26 to 30 June 2015, the worst flood in 600 years affected over 65,000 people in Changzhou City, resulting in 410 million RMB in property damage. Almost all of the city’s regions were directly affected, including five districts and two county-level cities (Figure 1). A new national precipitation record of 247 mm in 24 h was documented. Liyang county and Wujin district were reported as the worst-hit areas in this flood, but their sub-districts (or towns) had different response performance levels during and after the disaster [41,42]. It is therefore useful to examine the factors most affecting resilience enhancement. Many reasons make it possible to conduct an empirical resilience study in Changzhou. First, records from rain gauges and stream measurement stations can be added to a database and used for further hazard stress and damage modeling in the ArcGIS software (Environmental Systems Research Institute, Inc., Redlands, CA, USA). Second, since the Ministry of Environmental Protection of China requires the selected enterprises (e.g., local power plants, sewage plants and factories) to monitor the waste-water discharge and waste-gas emission at the drain/chimney outlets or boiler mouths and disclose the data online every hour, the refunction of these key utilities (power plants and sewage treatment plants) and factories can help monitor infrastructure recovery process and provide a reference when assessing community resilience to flood disasters. Third, the latest census was conducted in 2010 and there has been no significant boundary change between 2010 and 2015, so the 2010 census data will provide much relevant and detailed socio-economic data for further analysis.

Figure 1. Study area.
2.2. Measuring Recovery Capability

Recovery is defined here as the process of reconstructing study regions, enabling life and livelihood to return to pre-disaster states [11]. Recovery capability is therefore represented by the length of time required to return to the pre-disaster status. In our case, the waste-water and waste-gas discharge/emission data of key utilities and main factories or their surrounding factories and residential areas provide a basis to measure such a capacity. This is because the amount of waste water and gas produced by people’s living and local industrial production retains in a stable level in normal situations; however, during the flooding period, these factories or their served factories and residential areas are inundated and therefore the hourly amount of waste-water discharge and waste-gas emission would be disturbed, resulting in abnormal records (extremely high or low values). As a corollary, it is reasonable to capitalize on the waste-water and waste-gas discharge/emission data to assess the recovery capability, given the fact that such data can reflect whether people’s living and local industrial production are at normal or abnormal status.

To estimate the duration of a recovery process, a new approach is proposed in this study, which applies a time-series analysis of waste-water and waste-gas discharge/emission data that are disclosed online by key monitoring enterprises. There are 42 discharge/emission monitoring points from sample enterprises recording waste-water and waste-gas discharge/emission data every hour (Figure 2). Given that most of the discharge/emission records showed a disturbance during the flood-time and the enterprises attributed the disturbance to flood influence, it can be reasonably assumed that there was a change in the time series of waste-water and waste-gas discharge/emission data if the enterprise had been affected by flooding. And thus we applied change detection analysis using the R programming language to extract the beginning and end times of the change, and used the results to estimate the duration of the recovery process of the sample enterprise. The time taken for the enterprises to recover is used to represent the recovery duration of regions in the study area. The following section describes the procedures in detail.

To minimize the influence of autocorrelation—the cross-correlation of a signal with itself at different points in time—plots of autocorrelation function (ACFs) and partial autocorrelation
functions (PACFs) were first examined to detect the autocorrelation degree. The differences of the time series data of each sample were then calculated, and finally ACF and PACF were rechecked to control the degree of autocorrelation. The detailed procedures are given in Figure 3.

![Figure 3. Workflow of data preprocessing](Image)

**Figure 3.** Workflow of data preprocessing: (1) This process is used to eliminate interference from the tendency of plot series and to finally achieve a stationary series, which usually works after the first difference. If not, the difference calculation continues (second difference, third difference, etc.) until the plot series is stationary; (2) This process is used to eliminate the interference from periodicity of the plot series, which should be carried out after the “tendency” difference, to remove the effect of such a “tendency.” The aim is to achieve an independent series over time. It usually works after calculating the first difference with 1-cycle lags (e.g., circadian series: lags = 24 h).

Next, we need to select an optimal model for change detection procedure. The results of distribution tests show that most samples do not belong to any known distributions, so nonparametric approaches should be used to detect the change. Additionally, since this study attempts to detect time points at the beginning and end of changes, at least two change points need to be detected in the final results. Therefore, optimal models can only be selected from nonparametric alternatives for multiple change point analysis.

Among a set of nonparametric models for multiple change points detection, models used in the changepoint, cpm and ecp packages in R serve as widely-used models for change detection, with proven effectiveness and efficiency in detecting anomalies in bioinformatics [43], transportation [44], finance [45], ecology [46], etc. Each of the nonparametric approaches has its pros and cons [47–49]. The model selection should be based on model performance in specific case studies. We first extracted change points in the observed time series of the sample enterprises using all three approaches, and then compared the detection results with real situations to achieve the optimal model. After determining the optimal model, the detection results were further modified for a better estimation of recovery duration.

The first step is to extract possible change points related to the flooding case. According to the precipitation records of all rain gauge stations in the study area, the first record of 24-h precipitation that exceeded 50 mm was at 26 June 2015 10:00, while the last was at 30 June 2015 9:00. Therefore, we can assume the start point of a possible change will be located during that period, if there is in reality a change over the flooding period. Based on this assumption, only points extracted between 26 June 2015 10:00 and 30 June 2015 9:00 are retained to detect the start point, and the start point of the change caused by the flood damage can only be one of these points. To determine the end point of the change, as the recovery process will only begin after a disaster strikes, points detected around 30 June 2015 9:00 (before or after) are hypothesized as the alternatives for end point selection. To further identify which of these alternatives could be the start or the end point of the change, we constrained the time duration of changes with respect to the time series of the precipitation data. As sample enterprises will always be disturbed (changes start) during the downpour and recover (changes end) after it, the start and end time of a downpour can serve as constraints in determining the exact start and end points. The downpour here is defined according to the Chinese precipitation classification system, where the degree of rainfall reaches downpour level if over 24 h precipitation accumulation exceeds...
50 mm and tends to continue. After examining the start and end time of a downpour recorded by the nearest rain gauge station, the start point of the change in a sample monitoring point should be the first immediately after the time of the first downpour record of its nearest rain gauge station, and the end point of the change should be the first of the alternatives immediately after the time of the last downpour record, as observed by its nearest rain gauge station.

Next, the detection results must be validated, which is an important but always difficult procedure in change detection. In this study, the validation procedure helps compare the performances of three approaches, and determines which is optimal for change detection. As the sample monitoring point demonstrates change only when the sample area or the surrounding area is damaged by flooding, we examined the final change detection results of sample enterprises with respect to the flood-damaged area. For this purpose 100 m buffers for each sample monitoring point were created and used to examine whether they intersected with the damaged area. Damaged areas are defined according to the inundation map created by the ArcHydro toolbox in ArcGIS software, based on water level data of gauge stations (Figure 4). The more the buffers of samples intersect with the damaged area, the better the detection results. The detection accuracy is calculated following Equations (1) and (2), based on real situations (damage or non-damage) and detection results (change or non-change), as shown in Table 1.

\[
\begin{align*}
\text{Correct Rate} & = \frac{a + d}{a + b + c + d} \\
\text{Misdetection Rate} & = \frac{b + c}{a + b + c + d}
\end{align*}
\]

where

- \( r_{\text{correct}} \): Correct detection rate;
- \( r_{\text{misdetect}} \): Misdetection rate.

![Figure 4. Distribution of gauge stations.](image)

**Table 1.** Comparison between detection results and real damage conditions.

<table>
<thead>
<tr>
<th>Sample Cases</th>
<th>Change</th>
<th>Non-Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Damage</td>
<td>(a)</td>
<td>(b)</td>
</tr>
<tr>
<td>Non-Damage</td>
<td>(c)</td>
<td>(d)</td>
</tr>
</tbody>
</table>
After validation of the detection results, the best of the three approaches is already determined. However, as any of the three detection methods may have some misdetection cases, the result of the best approach should be modified to lower the misdetection rate. Misdetection can occur from two mismatching conditions: samples in the damaged area but with no changes detected (damage but non-change samples), and samples that have changes but are located in non-damaged regions (non-damage but change samples). For samples in the first condition, detection results of the other two approaches (mvc and cpm) should be checked, serving as substitutes for damage but non-change cases when similar detection results are found in both approaches. For samples in the second condition, changes detected by the selected method are compared to the results of the other approaches. When the other two methods have generated similar results, these common detection results will replace those detected by the selected method.

It is also noteworthy that changes detected in the previous process are composed of two parts—the duration of disturbance due to flooding damage, and the time consumption for a sample monitoring point to return to a stable status. To simplify the calculation procedure, the disturbance duration of a sample monitoring point is assumed to be the length of time the precipitation record of its nearest observatory remains in the downpour degree. Under this assumption, the returning time can be calculated by the following Equation (3).

\[
D_{\text{Return}}(i) = T_{\text{End of Change detected}}(i) - T_{\text{End of Downpour}}(i)
\]  

where

- \(D_{\text{Return}}(i)\): time consumption of \(i\)th sample monitoring point to return;
- \(T_{\text{End of Change detected}}(i)\): end time point of detected change of \(i\)th sample monitoring point;
- \(T_{\text{End of Downpour}}(i)\): end time point of downpour of \(i\)th sample monitoring point.

As sample enterprises are power plants, sewage treatment plants and main factories, their status can serve as an important index of production and life conditions in the neighboring area, and the return duration of these enterprises can to some extent represent the recovery capability of that area, particularly if it consists of build-up area which is derived from artificial land cover of the GLOBEINS30 dataset [50]. Therefore, to evaluate the recovery capability of the study area, we mapped the return time consumption over the build-up area, based on the calculation results of the sample enterprises.

2.3. Connecting the Measurements of Recovery to Resilience

The existing literature has shown that infrastructure recovery capability could serve as important proxy of community recovery capability [24,33–35,51] and provide a reference when assessing community resilience to natural disasters [52]. Therefore, in this step, the measured infrastructure recovery capability would be used as the external validation metrics to identify dominant resilience factors in the study area. In doing so, a number of variables were collected to represent the multi-dimensional nature of disaster resilience. The variables were selected from social, economic, infrastructural, and environmental dimensions, which are reportedly common components [53]. As several variables were from census statistics, such as demographic data at the sub-district (or town) level, other continuous variables were aggregated at the same levels to match the scale of the data. The final list of selected variables is given in Table 2. The first component encompasses social aspects, integrating age distribution, sex ratio, informal settler, and health service. Demographical attributes serving as subcomponents of the social component suggest that regions with fewer elderly people, children, or women and with a low level of migration are more likely to enhance community resilience, thus speeding up the recovery process after a natural disaster [54,55]. The variable of health service is represented by the share of health care facilities within a sub-district (or town), and a higher level suggests a higher standard of living for local residents, which may promote pre-disaster preparedness and post-disaster recovery [12,16]. The second component focuses on economic indicators, combining indices of industrial development and economic stability. Key indicators are the urban to rural ratio, GDP per capita, share of central business district (CBD), and manufacturing density. A higher ratio of urban to rural, and a greater GDP per capita will in
most cases indicate greater diversity and higher economic stability, which will enhance economic component [11,16]. The share of CBD and the density of manufacturing represent the commercial and manufacturing establishments that affect the economic asset exposure to natural hazards, implying a longer recovery time once damaged [11,16,56]. The third component summarizes infrastructure conditions; it is designed to evaluate the preparedness for resisting damage before a disaster, and to evaluate the rapidity and redundancy of response during and recovery after a disaster. Indicators of this component, such as access to administration centers, hospitals, and open spaces, provide an assessment of the capability of communities to deal with emergencies [55,57]. Road density is also included, to represent pre-disaster evacuation capability and the ability to efficiently respond and quickly recover after disasters [11,12]. Environmental resilience is the last component of disaster resilience, and variables collected include ecological conditions such as river density, urban green level, elevation, and slope. Previous literature had found that a larger amount of urban green areas can improve the ecological condition [15], while lower river density and a flatter surface reduce the risk of storm surge inundation and secondary landslides [12]. Lower elevation and slope also provide an improvement of the accessibility and ease of rescue work [13,58].

Table 2. Variables representing the four subcomponents of disaster resilience.

<table>
<thead>
<tr>
<th>Sub-Component</th>
<th>Label</th>
<th>Variable</th>
<th>Measures</th>
<th>Data Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Social Component</td>
<td>Age Distribution</td>
<td>Age</td>
<td>% population between 15 and 64 years old</td>
<td>PCC, 2010</td>
</tr>
<tr>
<td></td>
<td>Sex Ratio</td>
<td>Gender</td>
<td>sex ratio (male/female)</td>
<td>PCC, 2010</td>
</tr>
<tr>
<td></td>
<td>Health Services</td>
<td>Support of health facilities</td>
<td>population share of health facilities</td>
<td>PCC, 2010</td>
</tr>
<tr>
<td></td>
<td>Non-household</td>
<td>Mobility of labor force</td>
<td>% non-household population</td>
<td>PCC, 2010</td>
</tr>
<tr>
<td></td>
<td>Population</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Economic Component</td>
<td>Ratio of Urban to Rural</td>
<td>Urban population</td>
<td>ratio of residents to villages</td>
<td>OSM, 2013</td>
</tr>
<tr>
<td></td>
<td>GDP per Capita</td>
<td>Gross Domestic Product</td>
<td>Gross Domestic Product per capita</td>
<td>GCRD, 2010</td>
</tr>
<tr>
<td></td>
<td>Share of CBD</td>
<td>Commercial establishment</td>
<td>Share of business (CBD centers) within build-up area</td>
<td>OSM, 2013</td>
</tr>
<tr>
<td></td>
<td>Manufacturing Density</td>
<td>Manufacturing establishment</td>
<td>density of manufacturing within build-up area</td>
<td>OSM, 2013</td>
</tr>
<tr>
<td>Infrastructural Component</td>
<td>Road Density</td>
<td>Transport</td>
<td>road density</td>
<td>OSM, 2013</td>
</tr>
<tr>
<td></td>
<td>Access to Open Space</td>
<td>Emergency preparation</td>
<td>access to open space (parks, urban green space, stadium, parking area)</td>
<td>OSM, 2013</td>
</tr>
<tr>
<td></td>
<td>Access to Administration</td>
<td>Emergency preparation</td>
<td>access to governmental institutions (administration center, police station)</td>
<td>OSM, 2013</td>
</tr>
<tr>
<td></td>
<td>Access to Hospital</td>
<td>Emergency preparation</td>
<td>access to hospitals</td>
<td>OSM, 2013</td>
</tr>
<tr>
<td>Environmental Component</td>
<td>DEM</td>
<td>Geological condition</td>
<td>mean of Digital Elevation Model (DEM) data</td>
<td>GDEMV2, 2009</td>
</tr>
<tr>
<td></td>
<td>Slope</td>
<td>Geological condition</td>
<td>mean slope</td>
<td>GDEM SLOPE, 2009</td>
</tr>
<tr>
<td></td>
<td>River Density</td>
<td>Land use in natural terms</td>
<td>density of water network</td>
<td>OSM, 2013</td>
</tr>
<tr>
<td></td>
<td>Urban Green Area</td>
<td>Green area within build-up area</td>
<td>total area of green space per square kilometer of build-up area</td>
<td>OSM, 2013</td>
</tr>
</tbody>
</table>

Note: * Statistic unit: per sub-district (or town); The demographic data was derived from the published 2010 Population Census of China (PCC); The GDP data was obtained from a 1 km grid GDP data of China (2010) (Global Change Research Data Publishing and Repository, 2014) (GCRD); Information Points were downloaded from the Open Street Map (OSM); Build-up Area was derived from the artificial land cover in GLOBELAND30 dataset [50]; DEM and slope data were derived from the ASTER GDEM dataset published in 2009.

After determining the variables to represent disaster resilience, the map of time consumption for recovery was spatially intersected and aggregated at sub-district (or town) level to generate the corresponding validation metrics at the same scale. A mean value of recovery duration per sub-
district (or town), was then generated to represent the recovery capability of each sub-district (or town). A longer duration of recovery implies a lower recovery capability.

To validate the indicators of disaster resilience, regression analysis was applied between the external validation term, represented by the recovery capability of each sub-district (or town), and the variables collected as proxy for measuring disaster resilience. To prepare the independent variables, all data in the sub-district (or town) were standardized through a “Min-Max” conversion resulting in “0–1 range” rescaled variables. Variables which were interpreted as highly correlated (Spearman’s $R > 0.700$) were eliminated from further consideration to collinearity problems. The return speed of water level within 24 h after the strong rainstorm was included in the regression as the control variable to avoid the confounding effects associated arising from the intrinsic relationship between floods intensity and flood recovery. This control variable further allows the separation of recovery process after the flood recession from the total recovery process including the process of flooding receding and the process of refunctining of people’s lives and livelihoods as well. As the latter was documented being largely affected by preexisting community resilience level [11,23], it could, in turn, be more reasonable to be used as the external term to validate metrics of resilience.

As preliminary testing of variables showed a violation of linearity and normality assumptions concerning regression, ordinal logistic regression analysis was applied to assess the association among recovery capability, selected resilience variables and the control variable. Ordinal logistic regression is applied to deal with a dichotomous dependent variable, so allowing for more than two (ordered) response categories. Therefore, to prepare the dependent variable for logistic regression, fifty-eight sub-districts (or towns) within the study area were classified into four classes. These are coded from 1 to 4, where 4 indicates the sub-district could recover within 24 h, representing the highest recovery capability; 3 indicates a sub-district with the second-highest recovery capability (24 h $\leq$ recover time $< 48$ h); 2 represents the third rank of recovery capability (48 h $\leq$ recover time $< 72$ h); and 1 represents the last ranked class, which takes over 72 h to recover.

Unlike linear regression analysis, logistic regression models the log odds ratio of outcome as a linear combination of the predictor variables (Equation (4)). The outcome is the probability of one specific event occurring. In this study, the highest recovery class (Class 4) was assigned as the reference class, and an event was considered to occur when there is movement from one recovery class to the next.

$$
\log \left( \frac{P(Y = k_j)}{P(Y = K)} \right) = a_i + b_{i1}x_1 + b_{i2}x_2 + \cdots + b_{ij}x_j
$$

In which,
- $P(Y = k_j)$: The probability of observing the particular set of dependent variable values ($k_j$) that occurs in the sample;
- $P(Y = K)$: The probability of referent dependent variable value ($K$) that occurs in the sample;
- $a_i$: The multinomial logit estimate for observing $k_j$ occurring, relative to observing $K$ occurring in the sample, when the dependent variables in the model are evaluated at zero.
- $b_{ij}$: The coefficient between $j$th independent variable ($x_j$) and the natural log of the odds of the dependent variable equaling $k_i$, when the coefficient is usually estimated using maximum likelihood.

3. Results

3.1. Results of Recovery Capability Measurements

For further explanation, a plot series of one sample monitoring point (WWS3) is taken as an example (Figure 5). The three pictures in Figure 5a illustrate the series plot, ACFs, and PACFs of the original data, respectively. In most cases, the ACFs of a stationary sequence usually displays a sharp cutoff; however, those of the original series decay much more slowly (i.e., significant spikes at higher lags), demonstrating that the original series is not stationary, and to become stationary it requires an order of differencing. A differenced series will help us examine the periodicity of the sample dataset,
as the obvious tendency of the original series may sometimes overwhelm its periodicity. Thus, the ACFs and PACFs of differenced series have been explored and are given as the middle and bottom figures in Figure 5b. As the ACFs decay quickly and have no significant spikes at higher lags, it appears that this differenced series is a stationary series, and does not need “periodic difference” to remove its periodicity.

Figure 5. Series plot, ACFs and PACFs: (a) original sample data; (b) differenced series. The autocorrelations are significant for a large number of lags, but perhaps those at lag-2 and above are merely due to the propagation of the autocorrelation at lag-1. This can be confirmed by the PACF plot. Note that the PACF plot of original data has another significant spike at lag-3 in addition to lag-1, meaning that the higher-order autocorrelations cannot be easily explained by the lag-1 autocorrelation. Differencing should therefore be performed for further analysis.

Based on the preprocessed series of waste-water and waste-gas discharge/emission data, the selected three approaches were used to detect the change points, respectively. To compare results of the three methods, a flood inundation map was generated, for reference in further comparison analysis. The map, given in Figure 6, illustrates the spatially variable flood depth within the study area. The darker shades of red represent a high degree of damage, where the greatest inundation level was suffered. The seriously damaged area lies mainly within Jintan, whereas the administrative center (Zhonglou, Tianning) experienced the least flood inundation damage. This distribution information (the proportion of inundated build-up area, Figure A1) derived from the ArcHydro toolbox inundation map is consistent with the local government damage report (the proportion of affected population) (Figure A2).

A comparison of the results of the three change detection approaches according to the flood inundation map are listed in Tables 3 and 4, which indicate that the ECP approach has given a comparatively higher correct detection rate, and is the optimal method for change detection in this case.
Table 3. Detection results of the three methods.

<table>
<thead>
<tr>
<th>Method</th>
<th>Change</th>
<th>Non-Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mvc</td>
<td>Damage</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td>Non-Damage</td>
<td>3</td>
</tr>
<tr>
<td>Cpm</td>
<td>Damage</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td>Non-Damage</td>
<td>8</td>
</tr>
<tr>
<td>Ecp</td>
<td>Damage</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td>Non-Damage</td>
<td>5</td>
</tr>
</tbody>
</table>

Table 4. Accuracy of detection.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Mvc</th>
<th>Cpm</th>
<th>Ecp</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correct Detection Rate</td>
<td>0.63</td>
<td>0.65</td>
<td>0.72</td>
</tr>
<tr>
<td>Misdetection Rate</td>
<td>0.37</td>
<td>0.35</td>
<td>0.28</td>
</tr>
</tbody>
</table>

The time consumption of each sample monitoring point to recover was calculated according to change detection analysis. The recovery map of the whole study was then generated as Figure 7. The darker shades of red represent a comparatively longer time for regions such as Zhenglu, Hengliu, Xueyang, and Bieqiao to recover their pre-disaster status. These regions are mainly within the suburban area, which suffered different levels of flood inundation and damage, but which all experienced longer recovery times than other regions. Also of interest is the quick recovery process that occurs in regions within Jintan, as these regions suffered the highest inundation damage according to the flood inundation map (Figure 6).
3.2. Results of Regression Analysis

As the dependent variable, the recovery capability of each sub-district (or town) is classified into four, based on the average time consumption for recovery of artificial land area within the corresponding sub-district. The distribution of this dependent variable is mapped in Figure 8. To explain this spatial heterogeneity of recovery capability, which serves as an external validation term, an ordinal regression model was built to assess the contribution of an individual variable as a proxy for measuring disaster resilience. After eliminating highly correlated variables, nine variables were remained as possible explanatory variables in regression analysis. To construct the control variable, the return speed of water level was interpolated and aggregated into sub-district (or town) level. Then a mean value of the return speed per sub-district (or town), was generated to represent the recovery capability of each sub-district (or town) (Figure 9). During a stepwise ordinal regression analysis, six variables were significant according to the final results (Table 5). The “Chi-square” statistic of the model indicates that the regression model as a whole is statistically significant ($p = 0.001$). The “Estimate” in Table 5 is the ordered log-odds (logit) regression coefficient, which reveals the change in the recovery capability level when there is a one-unit increase in a potential independent variable, given that other variables are constant in the model. Coupled with “Significance”, these two parameters could help to identify more influential variables as proxies of resilience measurement to predict recovery capability.

In the social component, both of the two variables—health services and sex ratio prove significant in predicting recovery capability. Higher levels of health services and larger percentages of male population are important factors that make up the social resilience of communities. As health services can provide efficient rescue work when facing disruption and guarantee higher health condition of the community, easy access to health services will benefit the recovery process [53]. The sex ratio may also contribute to predicting recovery capability, as it is possible that males are in a
better position to respond to natural hazard events than females, and more able to participate in rebuilding work after natural disasters [59].

Table 5. Regression results of potential variables.

<table>
<thead>
<tr>
<th>Component</th>
<th>Variable</th>
<th>Estimate</th>
<th>Significance (Two-Tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Social</td>
<td>Sex Ratio</td>
<td>1.074</td>
<td>0.050 *</td>
</tr>
<tr>
<td></td>
<td>Health Service</td>
<td>1.219</td>
<td>0.043 *</td>
</tr>
<tr>
<td>Economic</td>
<td>Ratio of Urban to Rural</td>
<td>1.393</td>
<td>0.044 *</td>
</tr>
<tr>
<td></td>
<td>Share of CBD</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Infrastructural</td>
<td>Road Density</td>
<td>3.009</td>
<td>0.003 **</td>
</tr>
<tr>
<td></td>
<td>Access to Open Space</td>
<td>2.537</td>
<td>0.002 **</td>
</tr>
<tr>
<td></td>
<td>Slope</td>
<td>1.176</td>
<td>0.014 *</td>
</tr>
<tr>
<td>Environmental</td>
<td>River Density</td>
<td>0.934</td>
<td>0.092</td>
</tr>
<tr>
<td></td>
<td>Urban Green Area</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Control</td>
<td>Return Speed of Water Level</td>
<td>1.053</td>
<td>0.010 **</td>
</tr>
</tbody>
</table>

Notes: All data were standardized using “Z-score” conversion before regression; for the whole model, significance = 0.001, pseudo $R^2$ (Nagelkerke) = 0.614; Significance of the test of parallel = 0.804 > 0.005, which means the ordinal model is acceptable; * Significant at 0.05; ** Significant at 0.01.

Figure 8. Recovery Class of each sub-district (or town): Class 4—high recovery capability; Class 3—moderate to high recovery capability; Class 2—moderate to low recovery capability; Class 1—low recovery capability.
Based on the regression statistics, the ratio of urban to rural contributed significantly to explain variance of the recovery capability. The strong prediction power of the ratio of urban to rural suggested that a higher urbanization level could speed up recovery process. Given the high association between this urbanization variable and GDP, it also indicated that urbanization could directly enhance community resilience through generating greater economic diversity and larger amount of economic resources; and could indirectly promote resilience development through boosting the GDP growth and subsequently producing higher economic stability [11,16]. On the other hand, share of CBD was not included according to the stepwise regression procedure. As the share of CBD represents the local commercial establishment, a higher share of CBD implies higher economic asset exposure and might cause longer time to recover once seriously damaged [16,56]. However, according to the flood inundation map (Figure 6), regions around CBD in Changzhou (e.g., Zhoulou and Tianming) were not seriously affected by the flood hazards. Therefore, there is not enough evidence to analyze the contribution of economic differences between spatial units in this model.

Within the infrastructural component of resilience, two variables included, road density and access to open space, were both strong predictors of recovery capability. Road density directly relates to evacuation capabilities before disasters and redundancies within supply routes for response and recovery after disasters [8,60]. This fact is also in accordance with the high correlations among road density and another two accessibility variables—access to hospital and access to administration, which implies increasing accessibility to critical infrastructures could enhance resilience and speed up local recovery process. It is also noteworthy that access to open space is again validated as an important factor in building flood resilience. A possible explanation is that it can serve as floodable land, which could store floodwater and sediments without incurring further damage to constructions in the region [57]. In this way, constructions (power plants, sewage treatment and local factories) with less damage could be expected to return to their normal functions more quickly.

The topographic slope was the only environmental predictor for resilience. Considering the high correlation between DEM and slope, higher recovery capacity would be expected in flatter area
because lower elevation and slope could improve the accessibility and ease the response and rescue work [58]. On the other hand, lower river density can reduce the risk of storm surge inundation, and it would shorten the recovery duration once damaged [12]. However, the contribution of river density to recovery measured in this study was not statistically significant. A possible explanation is that river density could affect mostly the speed of flood recession but contribute less to the return speed of living activity and production. In the case of urban green area, the limited amount of green area over the whole build-up regions might constrain its contribution to recovery process.

The significance of the control variable highlights the influence of the return speed of water level in determining the total recovery capability. The inclusion of the control variable allows us to identify the contributions of other factors, especially to the refunctioning of people’s lives and livelihoods. As the recovery process of people’s living and productive activity was documented mainly related with antecedent resilience level, it is more reasonable and reliable to use the variables validated in this study as proxies of resilience to flooding disasters.

4. Conclusions and Discussion

Building flood resilience is a topic of considerable interest for governments, researchers, and residents prone to disaster damage, due to the increasing frequency of record-breaking flood levels in Asian delta cities. Within this context, a quantitative approach to assessing disaster resilience has previously been challenging. As there is no agreement on the precise definition of disaster resilience, it is hard to measure it directly. The traditional approach is to apply variables from multi-dimensions as proxy measures to index disaster resilience. To validate the selection of these variables, a method has been proposed that uses related empirical measurements (vulnerability, recovery, etc.) as external validation metrics. However, a shortage of available data, particularly data with high spatial and temporal resolution, constrains the application of this approach when measuring disaster resilience. This work aimed to address this issue.

This study has proposed a new method to measure recovery capability, using change detection analysis based on the time series of waste-water and waste-gas discharge/emission data. To guarantee the accuracy rate of change detection outcomes, the detection results of three approaches with respect to real observation data were compared, and ecp was selected as the final detection method. The change detection results were then modified according to the outcomes of the other two methods, to improve the accuracy rate. This approach made it possible to measure recovery after a flood hazard, and in most cases comparatively small-scale and short-term recovery processes were experienced. As the government of China has compelled local power plants, sewage treatments and industries (especially those with potential to emit heavy pollutants) to published their waste-water and waste-gas discharge data online, the method used here could have a broad application, especially in regions with limited access to other data sources to monitor recovery process (e.g., high-resolution remote sensing data).

To further explore variables that could be used to measure resilience, we applied ordinal regression analysis using the measured recovery capability at sub-district level, which could serve as an external validation term for selecting variables. The results show that health service, sex ratio, ratio of urban to rural, access to open space, road density and slope condition are statistically significant in characterizing disaster resilience. The six variables validated in this study are consistent with the set of variables mentioned as proxy for community resilience to other natural hazards [8,13,16,61]. This implies that, in general, investment in promoting adaptation activities and policies for one kind of natural hazard should have a synergistically positive effect on community resilience to other natural hazards; and to flash flood hazards, especially in regions within China, the six variables play more dominant roles in determining the recovery capability that was represented by the return speed of people’s living and industrial productivity. Therefore policies and measures related to the validated variables are suggested to be given priority when building up community resilience to flood hazards in the future. Specifically, a set of specific measures for the local government and urban planners are devised. First, based on the performance of the two socio-economic indicators which represent sex ratio and urban to rural ratio condition, respectively, it would be advisable for the local
government to reallocate more rescue resources to regions with higher proportion of vulnerable population (such as women, children, and the elderly) and to regions with lower level of urbanization. On the other hand, with respect to infrastructural and environmental aspects, decision makers and urban designers could suggest providing more open space (e.g., parks, stadium, and public squares) or enlarging existing ones; enhancing the accessibility and connections of a region where road network is undeveloped; setting up emergency services in residential areas near mountain regions or migrating residents from seriously vulnerable regions to more suitable areas.

As this work is the first attempt to use the proposed method in a real case study, there is still room for improvement along this line of research. For example, the explanatory power of the final model could be improved by introducing new independent variables, such as those relating to the effect of government policy and community capital on post-disaster recovery (e.g., the indicators related to emergency relief services, sense of place and cultural resources, etc.). If the data are available, these variables would be considered and their contributions in improving resilience explored in future studies. In addition, effective ways need to be examined to improve the accuracy of change detection and control uncertainties in validating change detection results. Besides the above aspects, other related empirical measurements can be considered (such as vulnerability, adaptive capacity and so on) in addition to recovery capability, to fully capture the characteristics of disaster resilience. And more empirical case studies at different regions may also be conducted to further evaluate the applicability of the proposed method.

Acknowledgment: We thank Isa Martinez Cano for her insightful comments on the revision of this paper. We would also like to thank the anonymous reviewers for valuable comments and suggestions. We acknowledge the waste water discharge and gas emission data from the key enterprises’ self-monitored pollutant discharge information platform of Jiangsu Province (http://218.94.78.61:8080/newPub/web/home.htm).

Author Contributions: All authors contributed to the concept, designed the study, discussed the results, and commented on the paper. B.H. and J.S. designed and performed the experiments. R.L. provided statistical expertise and revised the manuscript. J.S. managed the literature searches, performed the statistical analysis and wrote the first draft of the manuscript. All authors have read and approved the final manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1. Comparison between disaster resistance and disaster resilience.

<table>
<thead>
<tr>
<th>Concepts</th>
<th>Disaster Resistance</th>
<th>Disaster Resilience</th>
</tr>
</thead>
<tbody>
<tr>
<td>Definition</td>
<td>A measure of how much the system changes</td>
<td>The extent to which it can recover after the source of change is removed</td>
</tr>
<tr>
<td>Context</td>
<td>Resist the onset and impact of a disaster</td>
<td>Absorb the impact of a disaster</td>
</tr>
<tr>
<td></td>
<td>Continue to function at close to normal capacity and capability</td>
<td>Return to normal operations with the least possible delay and minimum possible dysfunction</td>
</tr>
<tr>
<td></td>
<td>Defeat the impact of the event</td>
<td>Limit the impact of the event</td>
</tr>
<tr>
<td>Measures</td>
<td>Mitigation</td>
<td>Preparedness</td>
</tr>
<tr>
<td></td>
<td>Preparedness</td>
<td>Response and recovery</td>
</tr>
<tr>
<td>Practical</td>
<td>Levees to deal with river flooding</td>
<td>Shelter and evacuation planning</td>
</tr>
<tr>
<td>Examples</td>
<td>Well-developed emergency services</td>
<td>Community disaster recovery task forces (e.g., post-disaster health services)</td>
</tr>
<tr>
<td>Pros and Cons</td>
<td>More specific and efficient measures to one type of disasters (e.g., Levees to flooding)</td>
<td>Measures to one type of disasters could be used for managing other disasters.</td>
</tr>
<tr>
<td></td>
<td>With limited utility against other threats</td>
<td>More complicated and need cooperation from different government departments, local institutions and agencies</td>
</tr>
<tr>
<td></td>
<td>Cost increases when taking specific resistance measures to each type of disasters</td>
<td>Cost decreases due to the common measures to different disasters.</td>
</tr>
<tr>
<td></td>
<td>Plan to achieve resistance where possible against high value threats</td>
<td>Achieve resilience for those we cannot reasonably resist, such as flooding caused by extreme precipitation</td>
</tr>
</tbody>
</table>


Figure A1. Inundated build-up area.

Figure A2. Comparison between the modeling results and local government damage report: each red bar represents the proportion of inundated build-up area (based on modeling results) and the blue line represents the proportion of affected population reported by the government in each district.

References


