Clustering for inundated meshes using tsunami simulation data and utilization of the clusters

*Sho Akagi[1], Tshihiko Hayakawa[1], Shin Aoi[2], Takahiro Maeda[2], Wataru Suzuki[2]

1. Mitsubishi Space Software Co., Ltd., 2. National Research Institute for Earth Science and Disaster Resilience

We have developed a database system called Tsunami Scenario Bank in which the water pressure fluctuation data at the offshore observation points (S-net) and the coastal tsunami height data and land inundation depth data in Kujukuri and Sotobo area are registered by carrying out numerical simulations of tsunami propagation and run-up for numerous source models that can cause tsunami damage in the coastal areas. We are also developing a system to forecast tsunami inundation from water pressure data observed offshore by selecting similar tsunami scenarios from Tsunami Scenario Bank (Multi-index method, Yamamoto et al., 2016). As in this system, the method selecting the scenario can predict the inundation distribution over a wide area at same time, on the other hand, it is difficult to verify the validity of the prediction in comparison with observation data at each land points or to understand the relation between selection of the tsunami scenarios and prediction inundation depth, in comparison with the method of predicting per point directly from observation data (Baba et al., 2004 etc.). In this research, we group meshes by using inundation depth data of Tsunami Scenario Bank to aim to understand nature of tsunami inundation distribution and form the basis for validation method to predicted inundation depth. For clustering, we adopted a combination of data handling for inundation depth and Principal Component Analysis (PCA) and K-Means method. First, in order to prevent unbalance of input data caused by a great difference in inundation depth depending on the source scale of the tsunami scenario, calculation to take logarithm to inundation depth of each mesh was applied, and the data of each mesh was reduced to the lower-dimensional subspace by Principal Component Analysis. Next, the K-Means method is applied and meshes at short distance in the subspace are grouped in a cluster. In other words, we classified meshes that have near inundation depth in the same scenario to be included in the same cluster. We show a result that about 1.85 million meshes in northern part of Kujukuri plain were grouped to 100 clusters by using the inundation depth data of 1,067 tsunami scenarios which set along the Japan Trench (Fig.1). The result of the clustering is based on the spatially continuous elevation, the distance from the river or the coastline, the positional relationship with the coastal protection facility. It is inferred that the tsunami inundation distribution takes these mesh clusters as spatial units. In addition, we investigated a method to predict inundation depth distribution in inland from inundation depth of clusters near coast where we can obtain tsunami run-up observations early, by using clusters relational analysis. For example, we show the result of estimation of the inundation depth in the inland clusters from the inundation depth of the coastal clusters, using the numerical simulation result of 2011 Earthquake off the Pacific Coast of Tohoku as imitation observation data (Fig.2). We found that this method can be expected as estimation of inundation depth distribution by simulation.

Keywords: Tsunami Simulation, Clustering, K-Means algorithm

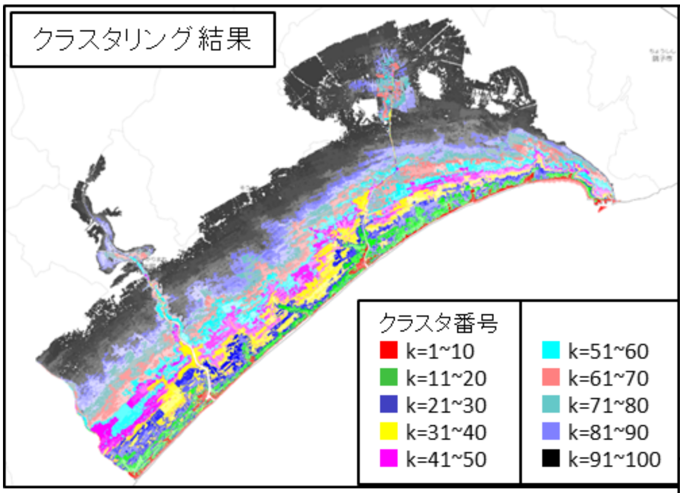

図1. 九十九里平野北部の計算メッシュ
クラスタリング結果

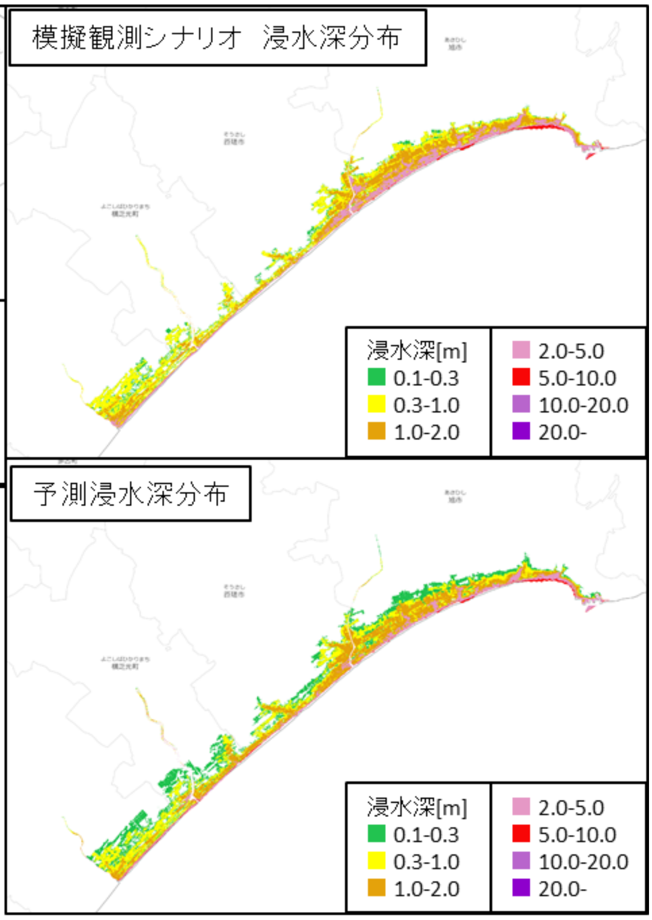

図2.九十九里平野北部
模擬観測シナリオ浸水深分布（上図）
予測浸水深分布（下図）