

—Data Paper—

Fine resolution basin database using 10 m DEM at the area along Sanriku coast of Japan

Takehisa Yamakita^{1,2*} and Hiroo Imaki³

This fine resolution basin database provides basin model over 10 ha sub-basins of the Sanriku coastal area located in the north west Pacific side of Japan. The data contain watershed boundary and flow path data about the target area created by use of a 10 m digital elevation model (DEM). Using this data, the coverage area was increased over 721.66 km² compared to the low resolution public data set provided by the government. It increased the coastline coverage from 17% to 84% compared with the low resolution data set in 1 km grid. For use in the coastal area or river discharge model, this data also contains the overall drainage basin boundary at the river mouth. The area of this survey was highly damaged by the tsunami on The 2011 Great East Japan Earthquake and post-disaster anthropogenic impact. Thus, this database will be of interest in the context of changes in land use and river discharge to assess both post-disaster anthropogenic impacts on the ecosystems.

Complete dataset is available via site: http://www.godac.jamstec.go.jp/catalog/data_catalog/metadataDisp/JAMSTEC-R_28DP01?lang=en&view=simple

Keywords : river basin, watershed, coast of Tohoku, 3.11 Great East Japan Earthquake, geographic information system (GIS)

Received 20 July 2017 ; Revised 10 August 2018 ; Accepted 11 September 2018

- 1 Project Team for Analyses of Changes in East Japan Marine Ecosystems,
Japan Agency for Marine-Earth Science and Technology (JAMSTEC)
- 2 Assessment of Environmental Dynamics Division, Department of Environmental Dynamics and Management, Graduate
School of Biosphere Science, Hiroshima University
- 3 Pacific Spatial Solutions, LLC (PSS)

*Corresponding author:

Takehisa Yamakita

Project Team for Analyses of Changes in East Japan Marine Ecosystems,
Japan Agency for Marine-Earth Science and Technology

2-15 Natsushima-cho, Yokosuka-city, Kanagawa 237-0061, Japan
yamakitat@jamstec.go.jp

Copyright by Japan Agency for Marine-Earth Science and Technology

1. Introduction

The 2011 Great East Japan Earthquake occurred on 11 March 2011. It caused huge tsunamis, long shaking and ground subsidence in the East zone of Japan and resulted in destruction of coastal areas especially in the area along the Pacific side of the Tohoku (North East) area of Japan. As well as to human society, the natural environment was severely damaged in many places (Yamakita et al., 2017a, 2018; Yamakita, 2018a). However, the consequence of tsunami is a natural disaster that has occurred repeatedly over a long time scale. Therefore, it is considered one of the factors characterizing the ecosystem in this area. On the other hand, construction, following the earthquake, to repair the sea walls and towns modifies the coastal environment and land use by artificially altering the environment. These changes are potentially new types of factors in the upheaval of the ecosystem that have not been seen in previous natural disasters. For this reason, consideration of broad scale human disturbance due to the earthquake and tsunami, is an important issue in the research on coastal ecosystems after the earthquake. However, such studies do not exist in broad spatial scales except for the flow of marine debris (Lebreton and Borrero, 2013).

In order to understand the impact of land-use changes on the watershed and coastal natural environment, it is common to assess river discharge effluents from changes in land use. A spatial model of the river flow is indispensable for this analysis. There are river basin model data sets available from the global level such as GDBD world basin database (Masutomi et al., 2009), Hydro SHEDS (Farr et al., 2007), and the national level, in the case of Japan. Use of such broad scale data has been successfully applied such as for the national goal of nature conservation, evaluation of ecosystem services or consideration of global water shortage (Oki and Kanae, 2006; Yamakita et al., 2017b; Yamakita, 2018a, 2018b).

However, there is a gap between the broad spatial scale management of basins using modelling with GIS databases and the local scale practical management using data of the individual rivers (Borowski and Hare, 2007; Yamazaki et al., 2017). This is caused by the resolution of data, differences in the types of data including data acquisition, and the purpose of the management. In the case of GDBD and Hydro SHED those are based on 1 km resolution data which is the derivative of GTOPO30 (USGS

<https://lta.cr.usgs.gov/GTOPO30>) and 3 arc-second data based on Shuttle Radar Topography Mission (SRTM). Using high-resolution and accurate basin data allows us to consider the potential application of broad-scale management (such as GIS modelling approaches) in local management.

Additionally, the Sanriku area, the region affected by the 2011 tsunami, shows complex coastline geomorphology so called Ria coast area. Thus, not many rivers can be covered by broad scale (i.e. lower resolution) data. Because of this situation, it is not possible to evaluate the impacts of terrestrial input in this area even if locally measured field data were available.

One of the ways to solve this problem is to have higher resolution basin data. In recent years, higher resolution digital elevation models (DEM) are more available worldwide. For example, the ASTER GDEM (ASTER-Global DEM) and AW3D-DEM (ALOS World 3D-DEM) has approximately 30 m resolution and coverage all over the world (Tachikawa et al., 2011; Tadono et al., 2015). However, relying on the single satellite product contain various observational errors (Yamazaki et al., 2017). In the case of Japan, 10 m resolution as public official base map which manually produced using aerial photographs and field survey are available from Geospatial Information Authority of Japan (GSI). In some locations 5 m resolution DEMs are also available, however; data coverage was not thorough at the time we downloaded data. Using these data it is possible to create accurate higher resolution basin models in the coastal area. Such basin data will be applied to the analysis of the effect of river input inside a bay and the effect of land modification on those relationships.

2. Method

2.1 Study site

We focused on the Pacific side of the Tohoku area, located in the temperate western North Pacific ecosystems which fall within one of the most productive area in the earth (Fig. 1). This area was heavily damaged by the tsunami on March 11, 2011 reached over 10 m and resulted in significant ground subsidence. The impact of the post-disaster anthropogenic activities, specifically in terms of rebuilding broad areas also resulted in major disturbances.

2.2 Used datasets and preprocessing for flow analysis

We obtained a 10 m digital elevation model (DEM) of this area from the digital elevation model of Base map information download service, provided by Geospatial Information Authority of Japan (GSI) (GSI, 2016). The final update was 10 May 2009. The data are the highest resolution covering the entire the study equally. Original data contains 0.1 m count elevation which guarantee 1 m accuracy extracted from standard topography map of 25,000 as 0.4 second grid point data. We used it under UTM N54 in 10 m resolution after applying open source DEM data transformation tool for Base map information elevation Ver1.6.6 by Ecolis Inc. which internally using GeoTransform function in GDAL (Ecolis, 2016). To model the stream flow, the sinks which is a cells without no lower neighbor has to be removed (general ideas of flow analysis was explained in the instruction of GDBD; NIES, 2009). To remove sinks caused by ponds or artifacts in the DEM, we used breaching sink removal methods (Jones, 2002; Lindsay, 2016b). This method requires less modification of DEM and more accurate than the simple filling method, which was commonly used in previous studies. Compare to the simple filling methods that increase elevation values of the cells lower than surrounding area, the breaching methods connect cells which have lower elevation values by degrading higher elevation cells in between.

We also used existing stream line data from National Land Numerical Information download service provided by Ministry of Land, Infrastructure and Transport

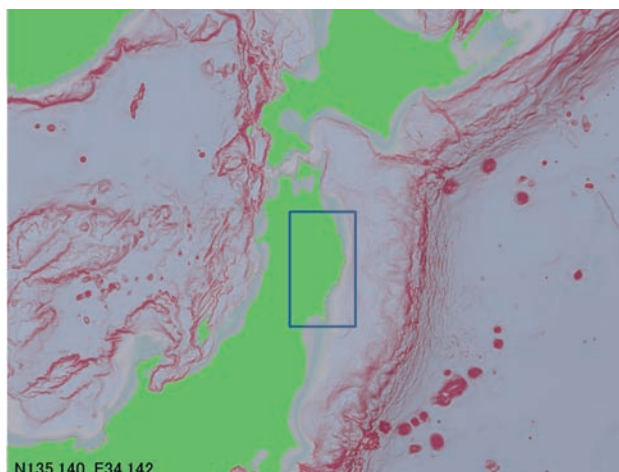


Fig. 1. Location of the study area, Sanriku, north east pacific side of Japan. Background image is CS stereograph of bathymetry data using JTOPO 30.

of Japan (MLIT, 2009) to condition DEM to enforce channel flows to match existing channels (burn in). To burn in the existing stream line decrement value, the distance decay value which defined by Lindsay (2016b) are specified as 100 and 0.5 successively. Thus the present stream lines have more likelihood to be stream line in the result, but when the calculation from new DEM exceed the likelihood with given weight new stream line will be selected. This framework will be the effective especially in the areas contain near flat area. We used an open-source desktop GIS and remote sensing software named Whitebox GAT for these preprocessing and latter calculations of the basins (Lindsay, 2016a, 2016c). Due to the limitation of memory we first obtained tentative watershed calculated in 40 m resolution resampled DEM and 10 km buffer from that watersheds are used for the calculation.

2.3 Creation of the stream line and boundary of basins

We created boundary of basins and flow lines based on the flow direction and flow accumulation analysis, calculated by the DEM. To calculate the flow direction (flow pointer), D8 algorithm, which assign channel flow directions based on the simple 8 azimuth angles, was used (Esri, 2014; Jenson and Domingue, 1988).

Using above DEM data flow accumulation raster and stream line raster were calculated. To define stream initiation points which is the most upstream point of streamline, we adapted 10 ha river sub-basin size as minimum size of the sub-basins based on visual comparison with topographic charts from the GSI.

As the river mouth points are used to build the basin boundary, we used all sink nodes (potential river mouth) calculated by the river link node algorithm in the Whitebox GAT. After the creation of river basin polygon, we also extracted area did not have river mouth in 10 ha smallest basin size limitation. It was extracted as Coastal Drainage Area (CDA), which did not have an exact river mouth.

2.4 Comparison with the dataset of broader scale

To evaluate the improved resolution of this basin model compared to the broader scale basin data provided by MLIT, we counted the overlap area of the basins. We also counted the coastline grids which did not have the river basin in 1 km resolution.

8282 coastal basins which cover 721.66 km² were add as new basins (Fig. 2). 3600 polygons which cover

146.45 km² were considered as the rest of the mixture of small basins. Average distances and standard deviation of these new basins from coast lines are $731.66 \text{ m} \pm 946.99$ and $32.71 \text{ m} \pm 58.63$ respectively. It was significantly close to the coast line compared to the distribution of sub-basins using the low resolution data. We eliminated one collapsed polygon in low resolution data by MLIT for this comparison.

Consistency with broader scale data provided by MLIT was evaluated with our finer scale data by calculating the overlapped area and mismatch of the boundary. (Out

of our 24156 km² basin area in our finer scale basins, 21225 km² overlapped with the broader scale basin data.) Among the overlapped area, 32295 area which showed under 0.01 km² area for each considered as mismatch of boundary which caused by the difference of resolution of the data or difference of mountain ridge in topography data. Although those were 41.09 km² area in total, it was only located in the boundary of the basins as showed by the red colored area in the Fig 2.

When we extract the basins within 1 km rasterized

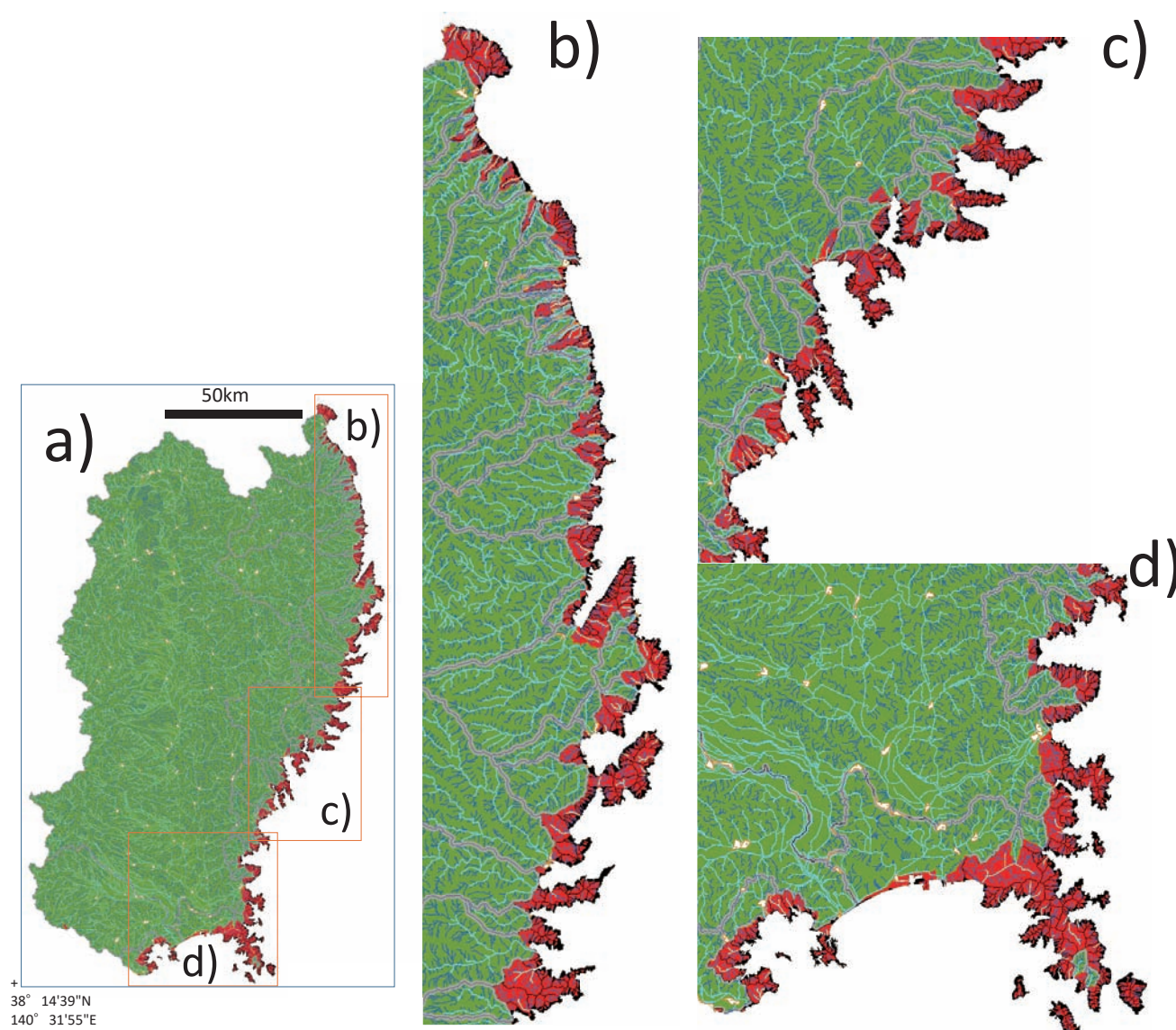


Fig. 2. Distribution of produced basins. Green area and right blue broad lines shows basins and stream lines in lower resolution using data obtained from the public data base. Red colored area, blue lines and black lines shows basins, stream lines and boundary of the basins, newly produced by our analysis. Black colored areas along the coast line show the remaining area which were not included in the basins. Mismatch with low resolution scale data showed by the yellow colored area and it was prevailed along the boundary of the basins from the public data base which showed by gray broad lines.

grids of the coastline, the number of grids which contain basins was increased from 165 to 626 grids. Although there are still rests of the mixture of small basins over 153 grids, our data of the river basins increased their coverage from 17.48% to 83.79% of the coast line.

3. Data format and contents

3.1 File format

The data files are saved in ESRI Shapefile as vector data format and GeoTiff raster image format which are commonly used for geographic information system (GIS) software. The shapefile consists of a main file, an index file, and a dBASE table with the extension of .shp, .shx and .dbf successively. We additionally provided map projection file and metadata file of those shapefiles in .prj and .xml format. The files are projected using Universal Transverse Mercator (UTM) projection and WGS 1984 UTM Zone 54 of northern hemisphere.

3.2 List of the files

The data provided as Shape geospatial vector data format for geographic information system (GIS) software (provided as zip compressed file). All data were coordinated in GCS WSG 1984 and projected as UTM Zone N54.

- tohoku002_100_05_framework.shp:

Polygon data of overall river basin and rest area (CDA).

- tohoku002_100_05_watershed.shp:

Polygon data of overall river basin over 10 km².

- tohoku002_100_05_subbasin.shp:

Polygon data of Sub-basin over 10 km². The data contains outside of our focused area until the 10 km buffer at the mountain side.

- tohoku002_100_05_stream_1000_selected_utm.shp:

Line data of the stream calculated by the model.

- tohoku002_100_05_stream_1000_sink.shp:

Point data of calculated river mouth. The data contains outside of our focused area until the 10 km buffer at the mountain side.

The data provided as GeoTiff raster image format.

- flow_accumulation.tif:

Flow accumulation raster in 10 m resolution. The value means the accumulated number of upstream cells.

- flow_direction.tif:

Flow direction in 10 m resolution. The value means the unique value of the compass direction (East = 1, South East = 2, South = 4, South West = 8, West = 16, North West = 32, North = 64, North East = 128).

3.3 Bounding box of each data

Extent of the data are as follows.

The data provided as Shape geospatial vector data format.

- tohoku002_100_05_framework.shp:

W 140.526264 E 142.082284

N 40.187951 S 38.240858

- tohoku002_100_05_watershed.shp:

W 140.526264 E 142.079877

N 40.187951 S 38.246872

- tohoku002_100_05_subbasin.shp:

W 140.409148 E 142.079880

N 40.188132 S 38.237680

- tohoku002_100_05_stream_1000_selected_utm.shp:

W 140.531252 E 142.079698

N 40.187366 S 38.247186

- tohoku002_100_05_stream_1000_sink.shp:

W 140.422167 E 142.069933

N 40.18536 S 38.245626

The data provided as GeoTiff raster image format.

- flow_accumulation.tif:

W 140.409113 E 142.180832

N 40.192186 S 38.202250.

Column 15082 Row 22021 Cell size 10 × 10 m

*Data under 0 m and area apart from 10 km buffer of our focus does not contain values.

- flow_direction.tif:

W 140.409113 E 142.180832

N 40.192186 S 38.202250

Column 15082 Row 22021 Cell size 10 × 10 m

*Data under 0 m and area apart from 10 km buffer of our focus does not contain values.

3.4 Details of the data table

Data table of each shape files are as follows.

The database file of the shape file named “tohoku002_100_05_framework.dbf” has latter parameter names:

- id: id number of the each area
- area_type: types of data (CDA or watershed)

The database file of the shape file named “tohoku002_100_05_watershed.dbf” has latter parameter names:

- area_type: types of data (all watershed)
- id: id number of the each area used in tohoku002_100_05_framework.dbf

The database file of the shape file named “tohoku002_100_05_subbasin.dbf” has latter parameter names:

- FID_1: id number of the each area

The database file of the shape file named “tohoku002_100_05_stream_1000_selected_utm.dbf”

has latter parameter names:

- FID_1: id number of the each line
- STRM_VA: types of data (all 1)
- Length: length of the each line in km

The database file of the shape file named “tohoku002_100_05_stream_1000_sink.dbf” has latter parameter names:

- FID_1: id number of the each line
- VALUE: types of data (all 5)

(Note: NA, not available.)

4. Expected use of the data

The database is expected to be used for research on the analysis of the effect of river input in coastal area, the effect of land modification on bays and so on.

5. Accessibility

This data is provided from Research Information and data Access Site of TEAMS (RIAS) which is the database of the JAMSTEC Tohoku Ecosystem-Associated Marine Sciences (TEAMS). <[http://www.i-teams.jp/catalog/rias/; \(ID:0202_R01\)>](http://www.i-teams.jp/catalog/rias/; (ID:0202_R01)>)

6. Usage Notes

Refer to this article for using our data. This is compatible with the Creative Commons Attribution 4.0 International license (CC-BY 4.0) (<https://creativecommons.org/licenses/by/4.0/>).

7. Ownership

These data belong to the Tohoku Ecosystem-Associated Marine Sciences (TEAMS) and Project Team for Analyses of Changes in East Japan Marine Ecosystems, JAMSTEC

Acknowledgements

We are grateful to Ms. Emily Egginton Skeehan for the English editing. We thank the editor and anonymous reviewers for providing valuable comments on the manuscript. This work was financially supported by the research program on Tohoku Ecosystem-Associated Marine Sciences (TEAMS) subsidized by the Ministry of Education, Culture, Sports, Science and Technology (MEXT).

References

- Borowski, I. and M. Hare (2007), Exploring the gap between water managers and researchers: Difficulties of model-based tools to support practical water management, *Water Resour. Manag.*, 21, 1049–1074, doi:10.1007/s11269-006-9098-z.
- Ecolis Inc. (2016), DEM data transformation tool for Base map information elevation Ver1.6.6, <<http://www.ecoris.co.jp/contents/demtool.html>>, (accessed 2016-11-01).
- Esri (2014), How Flow Accumulation works, <<http://resources.arcgis.com/en/help/main/10.2/009z/009z00000062000000.htm>>, (Last modified: 2014-04-10, accessed 2018-06-07).
- Farr, T.G., P.A. Rosen, E. Caro, R. Crippen, R. Duren, S. Hensley, M. Kobrick, M. Paller, E. Rodriguez, L. Roth, D. Seal, S. Shaffer, J. Shimada, J. Umland, M. Werner, M. Oskin, D. Burbank, and D. Alsdorf (2007), The shuttle radar topography mission, *Rev. Geophys.*, 45(2), doi:10.1029/2005RG000183.

- Geospatial Information Authority of Japan (GSI) (2016), Base map information download service, <<http://fgd.gsi.go.jp/download/>>, (accessed 2016-11-01).
- Jenson, S.K. and J.O. Domingue (1988), Extracting Topographic Structure from Digital Elevation Data for Geographic Information System Analysis, *Photogramm. Eng. Remote Sensing*, 54(11), 1593–1600.
- Jones, R. (2002), Algorithms for using a DEM for mapping catchment areas of stream sediment samples, *Comput. Geosci.*, 28, 1051–1060, doi:10.1016/S0098-3004(02)00022-5.
- Lebreton, L.C.M. and J.C. Borrero (2013), Modeling the transport and accumulation floating debris generated by the 11 March 2011 Tohoku tsunami, *Mar. Pollut. Bull.*, 66, 53–58.
- Lindsay, J.B. (2016a), Whitebox GAT: A case study in geomorphometric analysis, *Comput. Geosci.*, 95, 75–84, doi:10.1016/j.cageo.2016.07.003.
- Lindsay, J.B. (2016b), Efficient hybrid breaching-filling sink removal methods for flow path enforcement in digital elevation models, *Hydrol. Process.*, 30, 846–857, doi:10.1002/hyp.10648.
- Lindsay, J.B. (2016c), The practice of DEM stream burning revisited, *Earth Surf. Process. Landf.*, 41, 658–668, doi:10.1002/esp.3888.
- Masutomi, Y., Y. Inui, K. Takahashi, and Y. Matsuoka (2009), Development of highly accurate global polygonal drainage basin data, *Hydrol. Process.*, 23, 572–584, doi:10.1002/hyp.7186.
- Ministry of Land Infrastructure Transport and Tourism (MLIT) (2009), River data, National Land Numerical Information Download Service, <<http://nlftp.mlit.go.jp/ksj/gml/datalist/KsjTmplt-W05.html>>, (accessed 2016-11-01).
- National Institute for Environmental Studies (NIES) (2009), Center for Global Environmental Research (CGER), Global Drainage Basin Database (GDBD), <http://www.cger.nies.go.jp/db/gdbd/gdbd_index_e.html>, (accessed 2018-08-01).
- Oki, T. and S. Kanae (2006), Global Hydrological Cycles and World Water Resources, *Science*, 313, 1068–1072, doi:10.1126/science.1128845.
- Tachikawa, T., M. Hato, M. Kaku, and A. Iwasaki (2011), Characteristics of ASTER GDEM version 2, *International Geoscience and remote sensing symposium (IGARSS) IEEE international, 2011*, 3657–3660.
- Tadono, T., J. Takaku, K. Tsutsui, F. Oda, and H. Nagai (2015), Status of “ALOS World 3D (AW3D)” global DSM generation, *International Geoscience and Remote Sensing Symposium (IGARSS) IEEE international, 2015*, 3822–3825.
- Yamakita, T. (2018a), Cultural ecosystem services of temperate coastal areas of Japan: Uses of tidal flats, sandy beaches, and rocky - coral areas. *8th Asian Wetland Symposium 2017 PROCEEDINGS*, 290–293.
- Yamakita, T. (2018b), Change of the ocean after the Great East Japan Earthquake -Utilization of Geographic Information System and GIScience-, *E-Research in Species Biology 2*.
- Yamakita, T., Y. Matsuoka, and S. Iwasaki (2017a), Impact of the 2011 Tohoku Earthquake on the Use of Tidal flat: a Case Study in Inner Tokyo Bay, *J. Environ. Inf. Sci.*, 2017 (1), 25–36, doi:10.11492/ceispapersen.2017.1_25.
- Yamakita, T., K. Sudo, Y. Jintsu-Uchifune, H. Yamamoto, and Y. Shirayama (2017b), Identification of important marine areas using ecologically or biologically significant areas (EBSAs) criteria in the East to Southeast Asia region and comparison with existing registered areas for the purpose of conservation, *Mar. Policy*, 81, 273–284, doi:10.1016/j.marpol.2017.03.040.
- Yamakita, T., H. Yokooka, Y. Fujiwara, M. Kawato, S. Tsuchida, S. Ishibashi, T. Kurokawa, and K. Fujikura (2018), Image dataset of ophiuroid and other deep sea benthic organisms in 2015 extracted from the survey off Sanriku, Japan, by the research following the Great East Japan Earthquake 2011, *Ecol. Res.*, 33, 285, doi:10.1007/s11284-018-1571-7.
- Yamazaki, D., D. Ikeshima, R. Tawatari, T. Yamaguchi, F. O’Loughlin, J.C. Neal, C.C. Sampson, S. Kanae, and P.D. Bates (2017), A high-accuracy map of global terrain elevations, *Geophys. Res. Lett.*, 44, 5844–5853.