

# Multi-model ensemble estimation of volume transport through the straits of the East/Japan Sea

Sooyeon Han<sup>1</sup> · Naoki Hirose<sup>1,2</sup> · Norihisa Usui<sup>3</sup> · Yasumasa Miyazawa<sup>4</sup>

Received: 8 April 2015 / Accepted: 5 October 2015 / Published online: 1 December 2015  
© The Author(s) 2015. This article is published with open access at Springerlink.com

**Abstract** The volume transports measured at the Korea/Tsushima, Tsugaru, and Soya/La Perouse Straits remain quantitatively inconsistent. However, data assimilation models at least provide a self-consistent budget despite subtle differences among the models. This study examined the seasonal variation of the volume transport using the multiple linear regression and ridge regression of multi-model ensemble (MME) methods to estimate more accurately transport at these straits by using four different data assimilation models. The MME outperformed all of the single models by reducing uncertainties, especially the multicollinearity problem with the ridge regression. However, the regression constants turned out to be inconsistent with each other if the MME was applied separately for each strait. The MME for a connected system was thus performed to find common constants for these straits. The estimation of this MME was found to be similar to the MME result of sea level difference (SLD). The estimated mean transport (2.43 Sv) was smaller than the measurement

data at the Korea/Tsushima Strait, but the calibrated transport of the Tsugaru Strait (1.63 Sv) was larger than the observed data. The MME results of transport and SLD also suggested that the standard deviation (STD) of the Korea/Tsushima Strait is larger than the STD of the observation, whereas the estimated results were almost identical to that observed for the Tsugaru and Soya/La Perouse Straits. The similarity between MME results enhances the reliability of the present MME estimation.

**Keywords** Multi-model ensemble · Ridge regression · Volume transport · Tsushima Strait · Tsugaru Strait · Soya Strait

## 1 Introduction

The East/Japan Sea (hereafter, EJS) is a semi-enclosed deep marginal sea surrounded by Korea, Russia, and Japan. It is nearly isolated from the Pacific Ocean, except for the surface throughflow. The water balance of EJS is determined mainly by inflow and outflow through the Korea/Tsushima (KT), Tsugaru (TG), Soya/La Perouse (SP), Tatar, and Kanmon Straits connecting it to the East China Sea, Pacific Ocean, and Sea of Okhotsk. The Tsushima warm current inflow through the KT Strait mostly exits to the North Pacific through the TG Strait and partly to the Sea of Okhotsk through the SP and Tartar straits. The TG and SP Straits are shorter and narrower compared with the KT Strait. The Tartar Strait also communicates with the Okhotsk Sea and EJS, although the corresponding volume transport is negligible (0.01 Sv; Yanagi 2002). As the Kanmon Strait is very narrow and shallow, its volume transport is also negligible. Seasonally, the transports at the KT and SP Straits are typically large in summer–autumn and small in winter, whereas the seasonal variation of the TG

Responsible Editor: Pierre De Mey

This article is part of the Topical Collection on *Coastal Ocean Forecasting Science supported by the GODAE OceanView Coastal Oceans and Shelf Seas Task Team (COSS-TT)*

✉ Sooyeon Han  
han@riam.kyushu-u.ac.jp

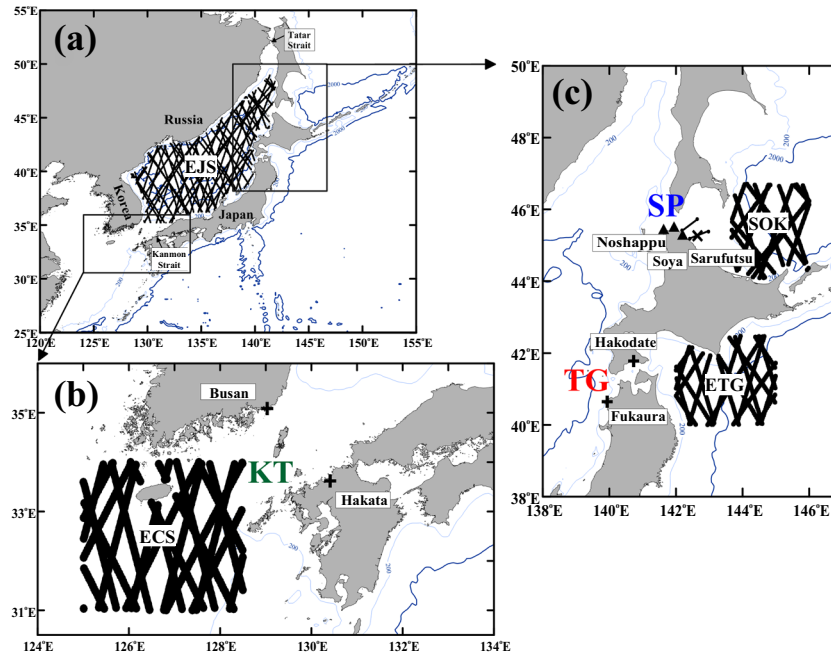
- <sup>1</sup> Interdisciplinary Graduate School of Engineering Sciences, Kyushu University, Fukuoka 816-8580, Japan
- <sup>2</sup> Research Institute for Applied Mechanics (RIAM), Kyushu University, Kasuga 816-8580, Japan
- <sup>3</sup> Oceanography and Geochemistry Research Department, Meteorological Research Institute, Tsukuba 305-0052, Japan
- <sup>4</sup> Application Laboratory, Japan Agency for Marine-Earth Science and Technology, Yokohama 236-0001, Japan

Strait is relatively small. The volume transport through the SP Strait has an annual range about twice that of the TG Strait, and the former has an annual mean volume transport of about half that of the latter (Seung et al. 2012).

The water mass entering the KT Strait should be basically balanced by the mass flowing out through the TG and SP Straits. However, this budget of observed data is inconsistent in the EJS system. Previous studies estimated seasonal and annual variations of volume transport through these straits using acoustic Doppler current profiler (ADCP) and high-frequency (HF) radar. The annual range of volume transport through the KT Strait varied widely from 0.5 to 4.2 Sv, and the seasonal variation of volume transport had a broad range of 0.7 to 4.6 Sv (Chang et al. 2004). The transport through the KT Strait has been measured by long-term ADCP using the ferry boat of the Research Institute for Applied Mechanics (RIAM) of Kyushu University since 1997 (Fig. 1). This estimation of the KT transport was about 2.6 Sv (Fukudome et al. 2010). Other studies have also estimated the volume transport through the KT Strait. Teague et al. (2002) proposed 2.7 Sv between May 1999 and March 2003 using 12 bottom-mounted ADCPs, and Isobe (1994) obtained an annually averaged volume transport of 2.3 Sv. The maximum transport reported by Fukudome et al. (2010) was greater by approximately 0.37 Sv, and the minimum transport was slightly greater than the estimation of Teague et al. (2002).

Previous studies suggested seasonal and interannual variations of transport through the TG Strait (Toba et al. 1982; Ito et al. 2003) with an annual mean of about 1.5 Sv. Nishida et al. (2003) also showed that the monthly mean transport through the TG Strait was 1.5 Sv by using ADCP only 22 times during 1993–1999. Ito et al. (2003) suggested that the volume transport from a direct and continuous measurement from November 1999 to March 2000 using the vessel-mounted ADCP decreased from 2.1 Sv on November 4 to 1.1 Sv on January 24 and March 15. The 4-month mean value was 1.5 Sv, which was in the same range as the previous studies.

Volume transport through the SP Strait was varied in the range of 0.5–1.5 Sv. The annual transport of the SP Strait was estimated to be 0.62–0.67 Sv from September 2006 to July 2008 and 0.94–1.04 Sv during 2004–2005 using bottom-mounted ADCP and HF radar. The difference between the two periods may be attributable to interannual variability of the SP current transport and/or the different measurement locations (Fukamachi et al. 2010). Matsuyama et al. (2006) suggested that the volume transport from other measurements was about 1.2–1.3 Sv in August 1998 and 1.5 Sv in July 2000. This discrepancy in the observed transports for the three straits may arise from a variety of sources such as observation periods, measurement devices, and exploiting methods from measurement data of ADCP or HF radar to volume transports.



**Fig. 1** Bathymetry showing the locations of the three straits of the East/Japan Sea (200m in light blue and 2000m in dark blue). Locations of AVISO along-track satellite altimeter data (black circle) available in the East/Japan Sea (EJS), East China Sea (ECS), Sea of Okhotsk regions (SOK) and east of Tsugaru Strait (ETG). KT, TG, and SP Straits refer to the Korea/Tsushima, Tsugaru, and Soya/La Perouse Straits, respectively. (a) Detailed information for each strait is shown in the enlarged maps (b)

and (c). (b) Locations (cross marks) of Busan and Hakata in a ferry boat (Camellia) line. (c) Locations of the tide-gauge station at Fukaura and Hakodate (cross marks) in TG Strait, and locations of the bottom-mounted ADCP measurement (northerly solid line), radar data grid points (triangles), ADCP location (cross), and radar section (southerly solid line) in SP Strait

On the other hand, ocean models at least provide a self-consistent budget despite subtle differences among the models. These estimations are able to simulate the volume transport of these straits. The amplitudes and phases of simulations are usually comparable to observations despite being virtual values. In recent years, ocean models including data assimilation have been estimated in this region by several organizations. These reanalyses can be more accurate than simulation without assimilation. Nevertheless, the estimated results have subtly different features depending on the model. We can present many reasons for the discrepancies among the models: the different physical process parameterization schemes, initial condition, and data assimilation.

Numerical model studies also suggested outflow partitioning in the EJS circulation using model results. Chu et al. (2001a) assumed that 75 % of the total inflow transport flows out of the EJS through the TG Strait and 25 % through the SP Strait. Bang et al. (1996) presumed to be similar in the ratio as 80 and 20 % respectively. Chu et al. (2001b) suggested that the ratio of the TG and SP transports is 0.61 and 0.39 using the US Navy Generalized Digital Environmental Model with variational P-vector method.

Furthermore, the observed transport through the TG Strait is about 70 % of the average of several estimates through the KT Strait. This ratio between volume transports through the KT and TG Straits was the first time estimated based upon concurrent observational data taken in the two straits (Na et al. 2009). The ratio between outgoing volume transports through the TG and SP Straits was 7:3, which was very close to the ratio suggested by Ohshima (1994), who applied the theory derived by Toulany and Garrett (1984) to understand the flow dynamics through the straits in the EJS. However, this ratio was based on relatively short observed time series.

To solve these discrepancies in observed data and to make optimal estimations, an approach that takes account of model uncertainties was needed; this approach has generally been considered to be the “multi-model ensemble.” The multi-model ensemble is generally easily identified as a simple multi-model ensemble. This simple multi-model ensemble is obtained by assigning equal weights to each of the models (Peng et al. 2002). To use this method, there should be a sufficiently large number of ensemble members to remove unexpected or unexplained ensembles, before the multi-model ensemble is conducted. However, it is difficult to abandon the poor ensembles given the small number of ensemble members or reanalyses of models. A more sophisticated approach for the multi-model ensemble seeks optimal multi-model ensemble predictions by obtaining different weights using multiple linear regression, a technique known as the multi-model superensemble developed by Krishnamurti et al. (1999a, b). The multi-model superensemble (for convenience, MME) approach was used in this study.

The MME has been used in the field of atmospheric science to examine uncertainties in models, but it has been underused in the field of oceanography to date. The following studies are based on its application to the atmospheric sciences. Krishnamurti et al. (2000) demonstrated that a multi-model ensemble outperforms all individual models for hurricane track and intensity forecasts. This multi-model ensemble was based on linear multiple regression of the different models against observations to determine statistical weights for each model. Hagedorn et al. (2005) also showed that the MME concept could improve single-model ensemble predictions and consistently estimate more accurately than estimation from any individual model.

Previous studies have indicated that the MME arising from a combination of multi-models with similar skill outperforms forecasts from individual models. Ideally, the models used should be as independent of each other as possible. As stated above, the volume transports of three straits from assimilated ocean models are similar despite subtle differences. The similar assimilation can also lead to a problem when at least one of the models is not entirely independent from the rest. That is, one of the input models in MME might include a certain small error by linear combination with the other ensemble members. This collinearity problem among the models is known as “multicollinearity.” Peña and Van den Dool (2008) assessed the performance of several consolidation methods that were divided into constrained and unconstrained multi-model ensemble forecast systems to predict monthly SST in the deep tropical Pacific. When multicollinearity existed in the models, ridge regression of the constrained consolidation methods was used to determine the optimal weight.

This study attempts to optimally combine the volume transports of the EJS system by using the MME approach. In addition to the MME approach, ridge regression was used to solve the multicollinearity among the assimilation models. The ridge regression approach has rarely been used in the ocean sciences.

The objectives of this study are twofold. First, we explore the importance of consolidation of four different data assimilation models in developing physical conservation transports in the EJS system. Here, we work to reduce the uncertainties of consolidation models applied for MME compared with the estimation of single models. Second, we compare various consolidation methods, particularly multiple linear regression and ridge regression. The multiple linear regression was based on the least squares method. The ridge regression is more complicated compared to the ordinary least squares method.

The sea level difference (SLD) not only across but also along a strait can be used to estimate the volume transport through the strait. The SLD with cross-channel is primarily in geostrophic balance, and the along SLD between the two oceans connected through the shallow strait is related with the hydraulic controlled (Garrett and Petrie 1981; Csanady 1982). In the KT Strait, Lyu and Kim (2003) showed that a strong

linear relationship exists between the transport and the SLD, using cross-strait hydrographic sections to remove baroclinic effects. Takikawa et al. (2005) demonstrated that the relations between the surface current velocities and the SLDs across the eastern and western channels in the KT Strait are approximately in geostrophic balance. The current entering the EJS may be regarded as being balanced with the outgoing transport to the Northwest Pacific. Previous studies have shown that the flow of the KT Strait is related to the SLD between EJS and the East China Sea (ECS) (Ohshima 1994; Toba et al. 1982). Additionally, the other straits are also linked to the SLD between the basin and the Pacific (Hata 1973; Ito et al. 2003; Ohshima 1994). According to Nishida et al. (2003), the volume transport of the TG Current is related to the SLD between Fukaura and Hakodate.

Considering these strong relationship between the volume transport and the SLDs, this study carries out the MME using these SLDs and the results compared with the MME result with transport. We examined the similarity between the two different estimation MME results to clarify the stability of solutions.

The article is organized as follows. The data used in the study are described in Section 2. Section 3 outlines the theoretical foundation of consolidation methods. Section 4 describes the MME results. Section 5 verifies the MME result, and Section 6 discusses and summarizes the results.

## 2 Description of data

The data used in this study consisted of volume transport and SLD data in three straits, the KT, TG, and SP Straits. The volume transport data comprised four different ocean models and observed data as ensembles to conduct the MME. Additionally, the SLD was considered as an independent variable for comparison with the MME result using transport data.

### 2.1 Measurement data

Measurement data have the role of explained variables in MME. As mentioned above, the observed data have inconsistency of budget, despite the fact that they were observed directly. Figure 2 summarizes the periods of the measurement

data. A 5-year overall period in the measurement data was determined based on the greatest overlap duration, and the volume transport was used to calculate a monthly average from January 2003 to December 2007.

In the KT Strait, the Research Institute for Applied Mechanics (RIAM) of Kyushu University has carried out long-term current measurements using a vessel-mounted ADCP between Hakata and Busan since February 1997 (Fukudome et al. 2010). The frequency of this ADCP data collection has been doubled from 6 or 7 to 12 or 14 times per week in accordance to the replacement of vessel in July 2004 (Fig. 2). However, there are several chances of errors in the observation data. The estimation of the volume transport ADCP measurement data has mechanical and process limitations. The ship-mounted ADCP is unable to measure the velocity near the bottom of the vessel. The data within the range of 15 % of the total depth from the seafloor also cannot be measured. Thus, the surface and bottom velocities are obtained by extrapolating the values at the shallowest and at the deepest depths of reliable measurements (Takikawa et al. 2005). The margin of error caused by these limitations maybe almost  $\pm 0.2$  Sv, assuming the error order of 0.1 m/s. In addition, the sampling intervals (the time between two successive cruises) vary from point to point through the ferry track. This problem may cause complicated tidal aliasing errors. Especially the  $S_1$  and  $S_2$  constituents possibly suffer from the infinite aliasing period at 12-hour measurement interval.

ADCP traverse observation cruise at the west mouth of the TG Strait was carried out seasonally for 22 times from 1993 to 1999. The relationship of the daily mean between volume transport and the SLD based upon the measured data is given by the linear equation by performing a regression analysis (Nishida et al. 2003)

$$Q = 0.0271 \Delta\eta + 0.933 \quad (1)$$

Where  $Q$  is the estimated volume transport of the TG Strait and  $\Delta\eta$  is the SLD between Fukaura and Hakodate (Fig. 1). In order to estimate the alternative transport, the TG transport is predicted by the regression model using the SLD data which are provided by the Japan Oceanographic Data Center (JODC), at the same tidal stations for the MME period from

**Fig. 2** Period and type of observed data in the major straits of the East/Japan Sea. The hatched area indicates a discontinuous period through the Tsugaru Strait

Names of strait	Period																				Data type
	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	
Korea / Tsushima																					vessel-mounted ADCP
Tsugaru																					Bottom-mounted ADCP Sea level
Soya / La Perouse																					Bottom-mounted ADCP and HF radar

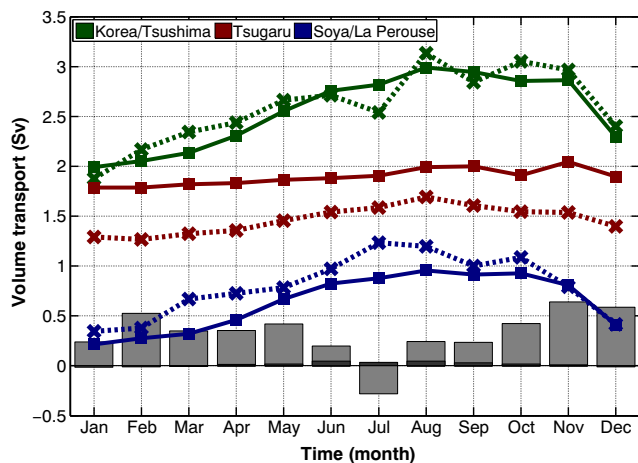
2003 to 2007. The predicted TG transport from the SLD includes considerable uncertainty, and the disagreement on the observed period between the regression analysis period and the predict period may also lead to substantial error.

The volume transport of the SP Strait was estimated during 2004 to 2008 using the combination of ADCP and HF radar data (Fukamachi et al. 2010) (Fig. 2). Compared with the data of the KT or TG Strait, the accuracy of data at the SP Strait may be the lowest since the HF radar system obtained only surface current information. Although the vertical structure was estimated with the assistance of ADCP data, the ADCP deployment site was downstream of the SP Strait and outside the ocean-radar coverage, and just one ADCP was deployed for 1 year.

The inconsistency of the observed data which is defined as the incoming transport minus outgoing one is 0.37 Sv at the annual mean. The water mass budget in the observed transports may be unbalanced due to the variety of error sources explained above. In addition, the unbalanced mass budget might be inevitable in non-synchronized observations. These are the reasons why we consider that the observed volume transports in these three straits of the EJS remain inconsistent (Fig. 3).

### 2.2 Multi-model description

The four different ocean data assimilation models used in this study, together with descriptions of their characteristics, are listed in Table 1. The reanalysis data from these models have represented similar patterns of seasonal variation with small differences, as we discuss later. The ensembles of the MME comprised four members: the Data assimilation Research of the East Asian Marine System (DREAMS), the



**Fig. 3** Monthly volume transports through the Korea/Tsushima (green), Tsugaru (red), and Soya/La Perouse (blue) Straits calculated from the observed data (cross) and DREAMS simulation (square). The bar at the bottom represents discrepancies between the transport entering the Korea/Tsushima Strait and transport outgoing through the Tsugaru and Soya/La Perouse Straits for observed data (light gray) and DREAMS reanalysis (dark gray)

Meteorological Research Institute Multivariate Ocean Variational Estimation System/Meteorological Research Institute Community Ocean Model (MOVE/MRI.COM, or for convenience MOVE), the Japan Coastal Ocean Predictability Experiment (JCOPE), and the HYbrid Coordinate Ocean Model (HYCOM). These members are from RIAM of Kyushu University, the Japan Meteorological Agency (JMA), the Japan Agency for Marine-Earth Science and Technology (JAMSTEC), and the Center for Ocean-Atmospheric Prediction Studies (COAPS) respectively. All models provide realistic data assimilation estimates. Table 1 summarizes the main model components and their data assimilation strategies. The Multi-variate Optimal Interpolation (MVOI) of data assimilation, which is taken into account in HYCOM, seems to be difficult to satisfy dynamical consistency such as the mass balance, considering the OI is more empirical method compared to near-optimal 3D, 4D-Var or KF. The period of analysis was 2003–2007, which was selected to match the configurations duration of the observed data (when all of the outputs coincide).

### 2.3 Sea-level data

The sea level data were used to examine the validity of the MME with volume transport data. The SLD data used consisted of two types, satellite altimeter data, and tide gauge data. The sea-level anomalies (SLA) measured by satellite altimeter, which were from Jason-1, Envisat, and GFO (plus available Topex/Poseidon and ERS-1/2 altimeters), were obtained from the Archiving, Validation and Interpretation of Satellite Oceanographic data (AVISO) (available at <http://www.aviso.oceanobs.com/>). AVISO has been distributing two types of altimeter data, near real-time data, and delayed-time data worldwide since 1992. The near real-time type data provide operational applications for directly usable high-quality altimeter data, but the delayed time products are more precise than the near real-time products due to their consistency. The spatial type is also divided into gridded and along-track products. The along-track product with delayed-time type data was used in this study.

The SLA used for 5 years from 2003 to 2007 were also averaged into monthly bins and then averaged in space. The monthly SLA of the EJS was calculated simply by averaging over the basin, with the exception of a 5-km band along the coastline. The averaged data represent an area that includes the ECS, east of TG (ETG), and the Sea of Okhotsk (SOK) regions, which are shaded in panels (b) and (c) of Fig. 1. The number of SLA data changed according to the time and spatial types. The number of data in each area is expressed as time series in Fig. 4. Data of the SOK region are not available during February–March 2003 due to the presence of sea ice.

To make the SLD through each channel from SLA, the SLD of the KT Strait was represented by the difference

**Table 1** Overview of models contributing to the multi-model ensemble

System name	DREAMS_B	MOVE/MRI.COM-WNP	JCOPE2	HYCOM+NCODA Global
Ocean model	RIAMOM	MRI.COM	Modified POMgcs	HYCOM
Domain	ECS+EJS	NW Pacific (around Japan)	Western North Pacific	Global
Horizontal Grid	1/4°×1/5°	1/10°×1/10°	1/12°×1/12°	1/12°×1/12° Orthogonal curvilinear
Vertical layers	z-coordinate 38 layers	Sigma-z hybrid coordinate 50 layers	Modified s-coordinate 45 layers	Hybrid coordinate (isopycnal/s/z) 32 layers
Nesting strategy	One-way nesting	One-way nesting	One-way nesting	One-way nesting
Atmospheric forcing	JRA-25 reanalysis; the GPV/GSM meteorological data	JMA's operational atmospheric analysis; Results of climate forecasting model	6-hourly NCEP Global reanalysis Forecast System or NCEP/NCAR reanalysis	NOGAPS; 3-hourly forcing QuikSCAT correction
Ocean data input	SSH	Jason+Envisat	NRL/SSC	Cooper-Haines projection
	SST	MGDSS	NAVOCEANO	Satellite
	In situ T,S	ARGO floats (T,S)	ARGO, Ship	XBTs, ARGO
Data assimilation scheme	RoKF	3DVAR with vertical coupled TS-EOF modes	3DVAR with vertical coupled TS-EOF modes	NCODA MVOI scheme
Agency / Institution	RIAM Kyushu University	JMA	JAMSTEC	Naval Research Laboratory

between the northeastern box of the ECS and EJS. Similarly, the other channels of EJS were given by the SLD between the EJS and the areas shown in Fig. 1c.

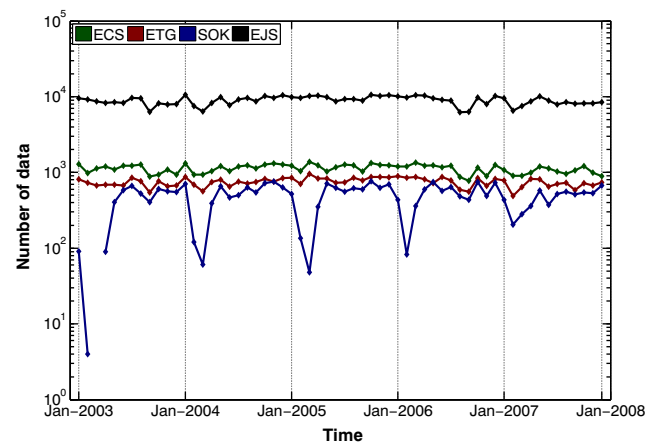
We also utilized tide gauge data obtained from the Japan Oceanographic Data Center (JODC) and the Korea Oceanographic Data Center (KODC). The data from the late 1900s are available at an hourly interval. The data used were taken from four points, which are marked in Fig. 1b, c, in the 8-year period from 2003 to 2010. The SLD between Busan and Hakata represents the volume transport of the KT Strait, and the SLD at Hakodate and Fukaura is related to the TG transport because the lines connecting the two points are nearly perpendicular to the axis of the currents at the straits. The SLD data for each strait, which were divided into the two periods 2003–2007 and 2008–2010, were also averaged by month. These monthly averaged data are used in prediction in Section 5.2.

### 3 Methods

#### 3.1 Multiple linear regression

A regression analysis is a statistical process for estimating the relationships between variables. Multiple linear regression (MLR), a term first used by Pearson (1908), attempts to describe the distribution of a dependent variable with the aid of a number of independent variables and to model the relationship between the dependent variable and one or more independent variables by fitting a linear equation to observed data.

The dependent variables are sometimes called regressands, or explained variables, whereas the independent variables are called regressors, or explanatory variables. Ideally, the models used should be as independent of each other as is possible so that their errors are small.



**Fig. 4** Numbers of monthly satellite altimeter data from January 2003 to December 2007 for the East China Sea (ECS), east of Tsugaru (ETG), Sea of Okhotsk (SOK), and East Japan Sea (EJS) regions, represented by green, red, blue, and black curves, respectively

The MLR is based on a least squares method, which means that the overall solution minimizes the sum of the squares of the differences between observed and predicted values from the results of each equation. The MLR expresses the value of a dependent variable as a linear function of one or more independent variables and an error term:

$$y(i) = \beta_0 + \beta_1x_1(i) + \beta_2x_2(i) + \dots + \beta_px_p(i) + \varepsilon(i) \quad (2)$$

where  $y(i)$  is the  $i$ th observation of the dependent variable and  $x_j(i)$  is the  $i$ th experiment on the  $j$ th independent variable,  $i=1, 2, \dots, n$  and  $j=1, 2, \dots, p$ . The values  $\beta_j$  represent parameters to be estimated, and  $\varepsilon(i)$  is the  $i$ th independent identically distributed normal error. Written over again in matrix form, one obtains

$$y = X\beta + \varepsilon \quad (3)$$

The method of ordinary least squares (OLS) for finding the values of the coefficients of the regression line is to minimize the sum of the squared vertical distance between the observed value and predicted value:

$$\min_{\hat{\beta}} \|\varepsilon\| = \min_{\hat{\beta}} \|y - X\hat{\beta}\| \quad \text{where } \|\varepsilon\| = \sum_i \varepsilon_i^2 \quad (4)$$

where solving for  $\beta$  yields

$$\hat{\beta} = (X^T X)^{-1} X^T Y \quad (5)$$

$\hat{\beta}$  is extrapolated under several assumptions. Above all of these assumptions, independent variables should not be correlated with each other. The models used should ideally be as independent of each other as possible so that their errors are small, although the monthly means of volume transport from the reanalyses are related to each other. In the KT Strait, the MOVE model is closely related to the DREAMS and JCOPE models, with correlation coefficients of 0.93 and 0.94, respectively. In addition to this, it is likely to be the same as that of the high correlation in other straits.

Similar models can also lead to problems when at least one of the models is not entirely independent of the others. The independent variables might be based on a presumption that one of the variables should be independent of the others. Alternatively, to say that two or more independent variables are independent means that the occurrence of one does not affect the probability of the others.

If collinearity exists among the independent variables, the result of the independent variables used is not appropriate for statistics. That is, one of the input models in MME might include a certain small error by a linear combination with the other ensemble members. This problem is known as “multicollinearity” in statistics. Multicollinearity refers to the presence of highly or

moderately intercorrelated predictor variables in ensemble members, and its effect is to invalidate some of the basic assumptions of the estimation of MLR.

To solve the harmful effects of the multicollinearity problem, spurious exogenous variables are dropped or ridge regression is used. It was difficult to drop spurious variables due to the paucity of ensemble members in this study, so another approach, the “ridge regression method” (Hoerl and Kennard 1970), was used to solve this effect.

### 3.2 Ridge regression

When multicollinearity exists in the model, it has several negative effects on the estimation result. First, the regression coefficients of individual models may change radically with the removal or addition of a predictor variable in the equation. Accordingly, the sequence of the weights can be switched. Second, the variance of the regression coefficients might be inflated even though the overall regression equation has good ability.

The variance inflation factor (VIF) is used as an indicator of multicollinearity in a matrix of predictor variables, that is, to determine how much the variance of an estimated regression coefficient is increased because of collinearity. Computationally, it is defined as

$$VIF_i = \frac{1}{1 - R_i^2} \quad (6)$$

where  $R_i^2$  is the coefficient of determination, which is a number that indicates how well the data fit a statistical model. Values of VIF that exceed 10 are often regarded as indicating multicollinearity, and values higher than 2.5 may be cause for concern.

When multicollinearity occurs, the  $X^T X$  matrix of the OLS estimator has a determinant that is close to zero, which makes it “ill-conditioned” so that the matrix cannot be inverted. If the OLS estimate was applied in the present condition in which ensemble members are correlated with each other, the estimates would be unbiased but their variances would be large, so the estimates may be far from the true value. There is the case, however, for which the “best linear unbiased estimator (BLUE)” is not necessarily the “best” estimator.

One approach to this is to use an estimator that is no longer unbiased, but has considerably less variance than the new least-squares estimator. A new way of doing this is ridge regression (RR), also known as Tikhonov regularization. RR seeks a solution for analyzing multiple regression data that suffer from multicollinearity and it is a multiple linear regression with an additional penalty term to constrain the size of the

squared weights in the minimization of the sum of the squared errors.

The RR estimate  $\hat{\beta}$  is defined as

$$\hat{\beta}_{\text{ridge}} = (X^T X + kI)^{-1} X^T Y, \quad k \geq 0 \quad (7)$$

where  $I$  denotes the identity matrix and  $k$  is a positive scalar parameter. A small positive value of  $k$  improves the conditioning of the problem and reduces the variance of the estimates. Although biased, the reduced variance of ridge estimates often results in a smaller mean square error when compared to least-squares estimates.

In this RR, the selection of  $k$  is important. Hoerl and Kennard (1970) proposed a method for selecting the correct value of  $k$ , which is the ridge trace by an iterative process. Typically,  $k$  begins with 0 and then runs through an increasing short interval. When the  $k$  value increases, the ridge coefficients begin tending toward zero, and a value is chosen when the ridge coefficients stabilize. Hoerl et al. (1975) attempted to determine the optimal value for  $k$  by use of the harmonic mean, and the solution is given by

$$k = \frac{p\sigma^2}{\hat{\beta}^T \hat{\beta}} \quad (8)$$

where  $p$  is the number of ensembles and  $\sigma^2$  is the residual mean square. However, when multicollinearity in the independent variables is extreme, i.e., independent variables are almost perfectly correlated, we would probably prefer to delete one or more independent variables before using the ridge approach as “stepwise regression.” However, the number of ensemble members in this study was only four, so this method was not available. The two consolidation methods, as stated above, were thus applied with the four different ocean models and the observed data to obtain more accurate transports in the EJS system, and the results of MLR and RR were evaluated for deterministic skill assessment.

## 4 Results

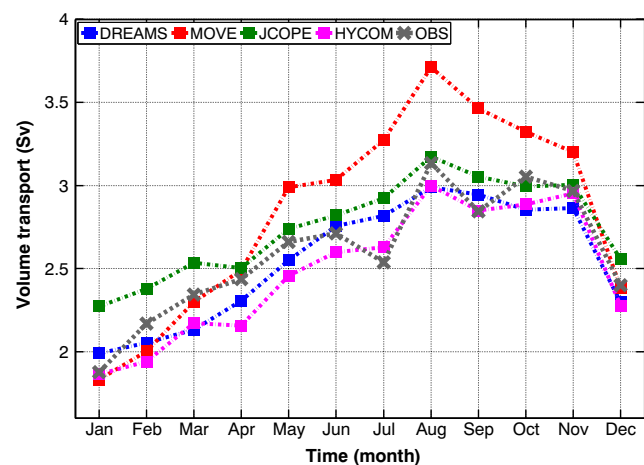
We attempted to conduct MMEs using the reanalyses from four different ocean models and the observed data to obtain more accurate data for volume transport at these straits. In Sections 4.1 and 4.2, we discuss in detail comparisons of results between individual and consolidation models, which are MLR and RR. Section 4.3 describes common coefficients to enhance physical conservation.

### 4.1 Comparison of single-model and multi-model ensemble

Figure 5 shows the seasonal variation of transports in the KT Strait using the reanalyses from the four single models and ADCP observation. The results of individual models have quantitatively different characteristics, although they show a similar tendency such as seasonal cycle. The MOVE reanalysis has the largest amplitude among the model results, which is almost twice larger variation range than the JCOPE output. The root mean square differences (RMSDs) between the observed data and the reanalyses of individual models vary widely in the range 0.10–0.28 Sv for the seasonal variation.

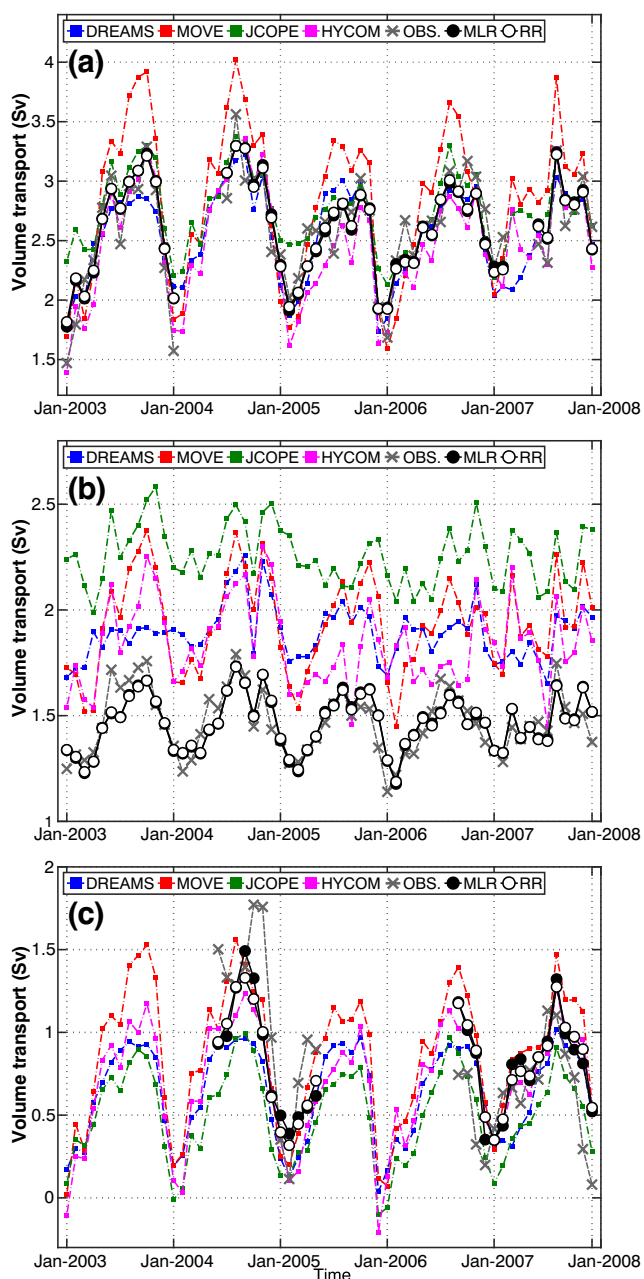
Comparisons of the monthly averaged transports through the KT, TG, and SP Straits from 2003 to 2007 between the observed data and the four model reanalyses are shown in Fig. 6a–c, respectively. The fluctuations of the observed volume transport in these straits are dominated by seasonal effects, and the range of seasonal variation differs from year to year. Each model reanalysis shows a similar tendency and also broadly resembles the observation. Nevertheless, some differences arise among the models, such as the amplitude and bias of transport. An apparent discrepancy is that all reanalyses from individual models for the TG Strait are overestimated in comparison with the observation (Fig. 6b).

Figure 7a–c shows the 5-year mean transports from 2003 to 2007 in these straits using the observed data and the model reanalyses. The 5-year mean transports of the MOVE reanalysis tend to be larger than the observation in all straits. The long-term mean transport of DREAMS has the smallest difference from the ADCP observations in the KT Strait. The 5-year mean of HYCOM, which also has a relatively small difference, well represents observation in the TG and SP Straits



**Fig. 5** Comparison of seasonal variation of volume transport from reanalysis from the four ocean models (square marks with colored curves) and ADCP observation (cross marks with gray curve) by month averaged from 2003 to 2007 through the Korea/Tsushima Strait





**Fig. 6** Monthly averaged transport through the Korea/Tsushima (a), Tsugaru (b), and Soya/La Perouse (c) Straits from January 2003 to December 2007. The observed data, reanalyses of models, and multi-model ensemble estimates are represented by the *gray cross mark*, *colored square mark*, and *black circle mark*, respectively. The consolidation models are consisted of multiple linear regression (*closed circle*) and ridge regression (*open circle*)

among the individual models. Particularly, overestimations of the modeled transports are manifested in the annual means through the TG Strait, with the largest gap of 0.79 Sv in the JCOPE reanalysis.

The MMEs using the observed data and four assimilation model results are expected to decrease these discrepancies. The 5-year mean transports of the consolidation models are similar to the measured transport, although all reanalyses from

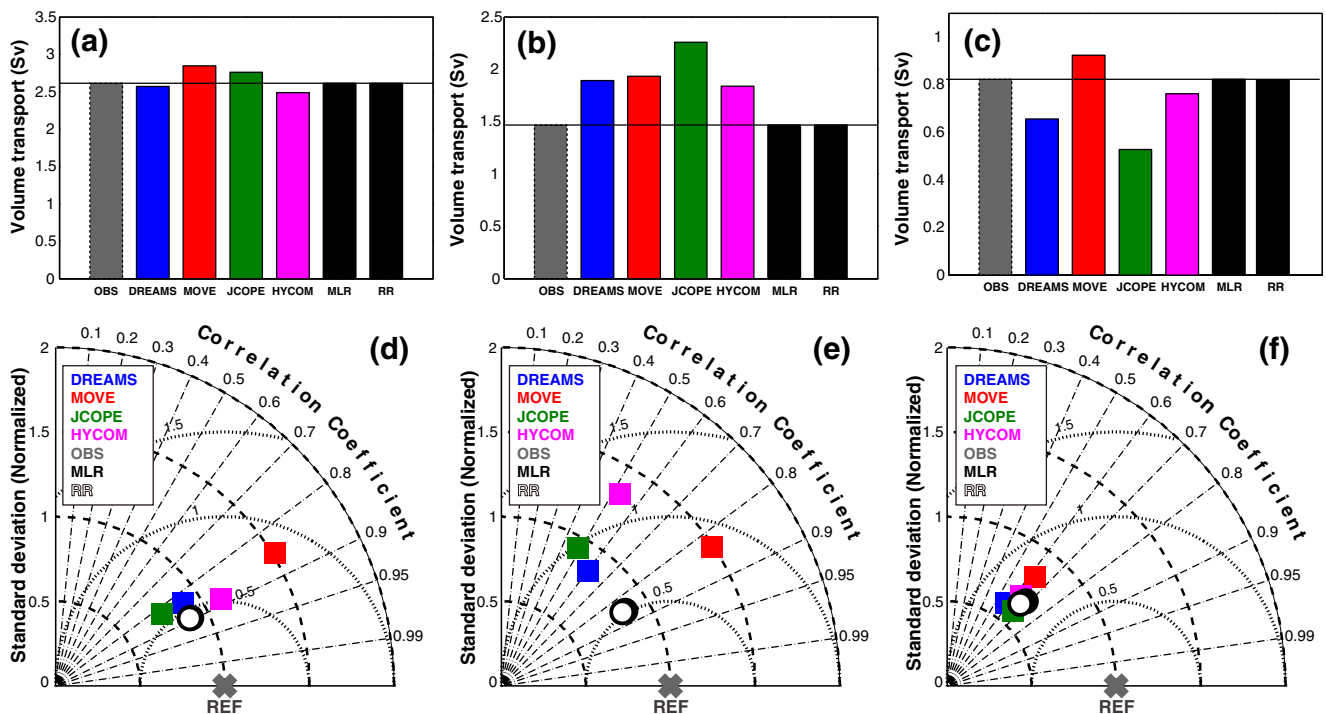
the single models strongly overestimated the volume transport in the TG Strait (Fig. 7b). The time-mean of the MME estimates can be calibrated to the dependent variable despite the fact that all independent variables are strongly biased. This suggests that the MME can be much closer to the dependent variable with the combination of “poor” ensembles.

To provide a detailed comparison of the single model reanalyses and the MME estimates, three important statistics are represented by diagrams developed by Taylor (2000) in Fig. 7d–f. The Taylor diagram enables visualization of the standard deviation (STD) of the model and observation patterns ( $\sigma_m, \sigma_o$ ), the correlation coefficient ( $R$ ), and the RMSD ( $E$ ) between the two fields simultaneously in a two-dimensional space by a polar coordinate system. These statistical measures are normalized to the observed STD. The normalized STD and squared difference can be written as:

$$\hat{\sigma} = \sigma_m / \sigma_o \tag{9}$$

$$\hat{E}^2 = \left( E / \sigma_o \right)^2 \tag{10}$$

We understand easily that each ensemble point quantifies how closely related the modeled field and observed field (represented as “reference” field) are on the basis of the three normalized statistics. The cosine of the angle of the model point from the horizontal axis of the Taylor diagram indicates the correlation between the observation and the model. The correlation coefficients between the observed data and the MME estimates are higher than those between the observed data and the individual models. For instance, the correlation coefficients between the observed data and the results of the individual models (DREAMS, MOVE, JCOPE, and HYCOM) for the TG Strait are 0.602, 0.835, 0.487, and 0.525, respectively (Fig. 7e), but the consolidation models have higher correlations with the observed data of about 0.856. The radial distances from the origin (0, 0) to the ensemble points in the Taylor diagram are proportional to the normalized STD. The STDs ( $\hat{\sigma}$ ) of the MME estimates are close to unity, similar to the normalized observed data, although each model has a standoff point from the observation in the TG Strait. This indicates that the MME is able to estimate the anomaly component of the observation data, although the MOVE and HYCOM points stray significantly from unity. Thus, the MME can estimate not only time-mean transport but also the anomaly component. The linear distance between the reference lying on the horizontal axis, and the point of the independent variable in the Taylor diagram is proportional to the RMSD. In the TG Strait, all models were remote from the observation, but the MME result is close to the reference (1, 0). In general, the normalized statistics of the MME estimates display



**Fig. 7** (upper panel) Five-year mean of volume transport through the Korea/Tsushima (a), Tsugaru (b), and Soya/La Perouse (c) Straits. The line is same as the observation value. (lower panel) Statistical analysis using a Taylor diagram for the volume transports of the Korea/Tsushima (d), Tsugaru (e), and Soya/La Perouse (f) Straits. The reference at the

bottom indicates the observation in each strait. The radial distance from the origin is proportional to the normalized standard deviation. The RMS difference between the model and reference field is proportional to their centered distance apart. The correlation between the two fields is given by the azimuthal position of the model field

the variability of volume transport more realistically than the individual models, although all ensemble members are not very close to the reference. The MME estimates are closer to the reference than any individual model. These statistics also expose thoughtful regression coefficients, as we discuss later.

The MME estimates are not very different for the other two straits. For the KT Strait, the normalized RMSDs ( $\hat{E}$ ) of the MME estimates are slightly smaller than those of any modeled point, and these RMSDs are pretty similar to the RMSD of HYCOM (Fig. 7d). It is important to note that the present MME analyses successfully eliminate the outlier effect of the MOVE, resulting in the minimum RMSD for the KT Strait.

However, the normalized RMSDs ( $\hat{E}$ ) and STDs ( $\hat{\sigma}$ ) of all of the single models and the MME estimates are almost identical in the SP Strait. The normalized STDs of the reanalyses from the models range from 0.585 to 0.828. After the MME was performed, the STDs of the consolidation models were about 0.665. The STDs of the estimates for MME remain underestimated in the SP Strait. This is considered to be a limitation of the MME, which is an unsatisfactory result, because all estimates of the individual models are too similar in the Taylor diagram (Fig. 7f). The inconclusive result is caused by collinearity among the independent variables, as we discuss in the next section.

## 4.2 Multiple linear regression and ridge regression

The volume transports of the SP Strait from the assimilated ocean models were similar (Fig. 7f). First of all, we examined how much multicollinearity exists within the independent variables. The VIFs were calculated to check the multicollinearity among the independent variables. The VIF of MOVE tended to be high in all straits (Table 2), which means that this ensemble is widely correlated with the other ensembles. The reanalyses of all models indicate strong VIF in the SP Strait but weak VIF in the TG Strait. In other words, the reanalyses from all models have many similarities, such as the variability and tendencies of the transport in the SP Strait, but these are independent of each other in the TG Strait.

Table 3 shows the equations obtained from the MME analyses for the MLR and RR for each strait. If the regression coefficients of each consolidation model are arranged in

**Table 2** VIFs of individual model transports through the Korea/Tsushima, Tsugaru, and Soya/La Perouse Straits

The predictors	DREAMS	MOVE	JCOPE	HYCOM
Korea/Tsushima	8.146	13.755	9.046	6.963
Tsugaru	1.740	2.627	2.744	3.249
Soya/La Perouse	11.378	19.523	17.131	16.955

**Table 3** Regression coefficients of multi-model ensembles for the Korea/Tsushima, Tsugaru, and Soya/La Perouse Straits

$y$	=	$\beta_0$	$+ \beta_1 x_1$ DREAMS	$+ \beta_2 x_2$ MOVE	$+ \beta_3 x_3$ JCOPE	$+ \beta_4 x_4$ HYCOM	TOTAL ( $\beta_1 + \beta_2 + \beta_3 + \beta_4$ )
Multiple linear regression							
$y_{\text{Korea/Tsushima}}$	=	0.546	+0.109	+0.147	+0.021	+0.527	0.804
$y_{\text{Tsugaru}}$	=	0.327	+0.211 $x_1$	+0.604 $x_2$	-0.087 $x_3$	-0.126 $x_4$	0.602
$y_{\text{Soya/La Perouse}}$	=	0.290	-0.849	-0.406	+1.544	+0.853	1.142
Ridge regression							
$y_{\text{Korea/Tsushima}}$	=	0.429	+0.153	+0.144	+0.114	+0.427	0.840
$y_{\text{Tsugaru}}$	=	0.283	+0.213 $x_1$	+0.567 $x_2$	-0.052 $x_3$	-0.108 $x_4$	0.620
$y_{\text{Soya/La Perouse}}$	=	0.189	-0.182	+0.099	+0.758	+0.341	1.016

descending order, the corresponding sequences of both the MLR and RR are almost identical (with the exception of the regression coefficients of JCOPE and HYCOM for the SP Strait). For the KT Strait, the regression coefficients of HYCOM are the largest and those of JCOPE are the smallest among the single models for both consolidation models. The similarities in the sequence of the regression coefficients between the two combination models indicate that both MME results are reliable for statistics. The residual terms,  $\epsilon$ , in the equations are almost zero in the obtained regressions for both MLR and RR (not shown in Table 3). The small difference between the observed and the MME estimates means that the MME is able to estimate the observed data.

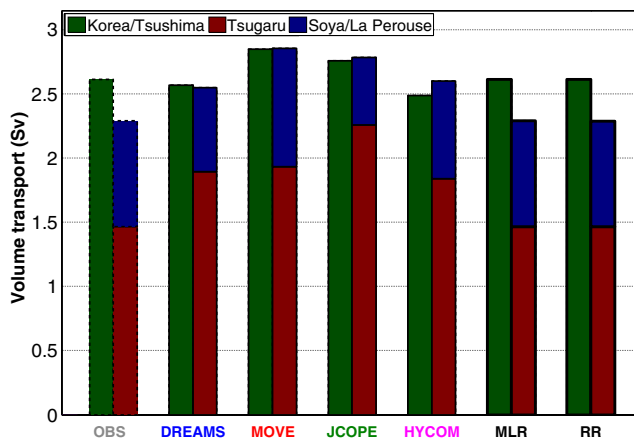
Many similarities exist, but we also find some discrepancies between the two MME estimates. The regression coefficients are quantitatively different depending on the consolidation method. Furthermore, the correlation coefficients between the observation and the MLR estimates (0.894, 0.856, and 0.682) are slightly higher than those of the RR estimates (0.893, 0.855, and 0.666) for the KT, TG, and SP Straits, respectively. The error terms  $\epsilon$  of the MLR estimates are also slightly smaller than those of the RR estimates. Although it seems that the MLR estimates outperformed the RR results, this assumption is incorrect. If the OLS estimator finds the regression coefficient of one of the independent variables, the others are considered constant numbers. In other words, the OLS estimator finds the estimation with minimized differences between the observation and the corresponding variable without considering collinearity within the independent variables.

There is evidence that the RR considers multicollinearity in the models in the comparison with the MLR. The most remarkable improvement is the relaxing of abnormal weights. A significant difference between the MLR and RR estimates is seen in the SP Strait. This difference comes from the multicollinearity of the dependent variables. It is exposed in the VIF for the SP Strait, which was larger than those of the other straits. On the other hand, both the MLR and RR results are almost identical for the TG Strait. These results are also related to the smallest VIF for the TG Strait. In a similar vein,

the decrease in the variation of the regression coefficients is greater when conducting RR than when conducting MLR for all straits. The intercept coefficient  $\beta_0$  in the regression equations indicates the amount by which the observation differs from the simulated fields. In other words, it also contributes to the bias of the MME model. If the independent variables are independent of each other and at least one of the explanatory variables can explain the explained variable, the intercept coefficient should be close to zero in the MME. The intercept coefficients of the RR analysis are smaller than those of the MLR for all straits, although all intercept coefficients are far from zero (Table 3). Although the applied results were almost identical as shown in Figs. 6 and 7, the present analysis demonstrates that the ridge estimator finds proper regression coefficients by relaxing multicollinearity problems compared with the least squares estimator. For these reasons, it is clear that the RR estimates outperform the MLR estimates.

However, some unsolved problems remain. The intercept coefficients  $\beta_0$  in the regression equations are not close to zero but have significant positive values (Table 3). This means that all single models include errors as large as  $\beta_0$  and/or that the observed data as the dependent variable have the problem of inconsistency of the mass balance. Figure 8 shows the differences between the transport entering the EJS and the transport flowing out through the TG and SP Straits from the observation data, the four model reanalyses, and the MME estimates. Neither of the MME estimates solves the inconsistency problem of the unbalanced budget, which is in common with the observation. Although the single-model ensembles at least provide a self-consistent budget, the MMEs do not satisfy physical conservation due to incomplete combinations of the multiple models. These problems are exposed by the total of the weights in the regression equations (Table 3). In particular, for the TG Strait, the total of the weights is relatively small due to the overestimation of the four models in the TG Strait.

According to these results, the estimates of the MMEs performed well statistically, but they did not satisfy the physical

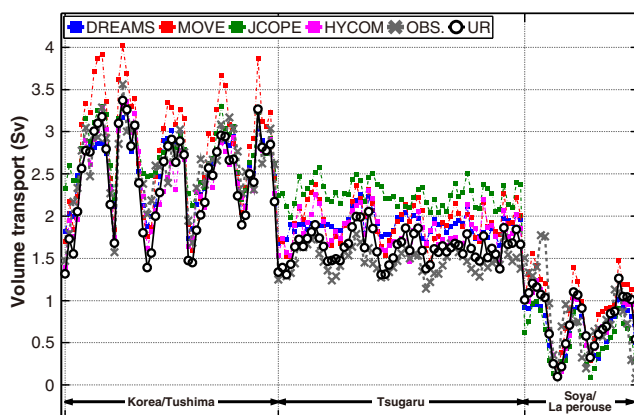


**Fig. 8** Differences between the volume transport entering the East/Japan Sea (green) and the outgoing transport through the Tsugaru (blue) and Soya/La Perouse (red) Straits for observed data, reanalyses from the four models, and consolidated estimations

conservation of volume transport. In addition, the estimated coefficients are scattered among the solutions, indicating the need for unified weights.

### 4.3 Unified estimation

To unify the coefficients for all of the straits, the MME was conducted for a connected system without dividing the three straits. Thus, the ensemble size was increased by about three-fold (Fig. 9). Although the number of ensembles is ideally 180 (monthly mean for 5 years for the three straits), the number of ensemble members used was actually 140 because of missing data for the SP Strait (Fig. 2). The present MME was carried out based on the same method as the previous MME. As stated above, the RR estimates outperformed the MLR estimates in solving the multicollinearity problem. Hereafter, we perform



**Fig. 9** Monthly averaged transport through the major straits. The  $x$ -axis includes each strait and period. The period differs according to the observed data of each strait. The observed data, reanalyses of the models, and ridge regression performance are represented by the gray cross mark, colored square mark, and black circle mark, respectively

only RR analysis on the connected system and give the new term “UR (Ridge regression applied to the unified system)” to distinguish from previous RR estimates.

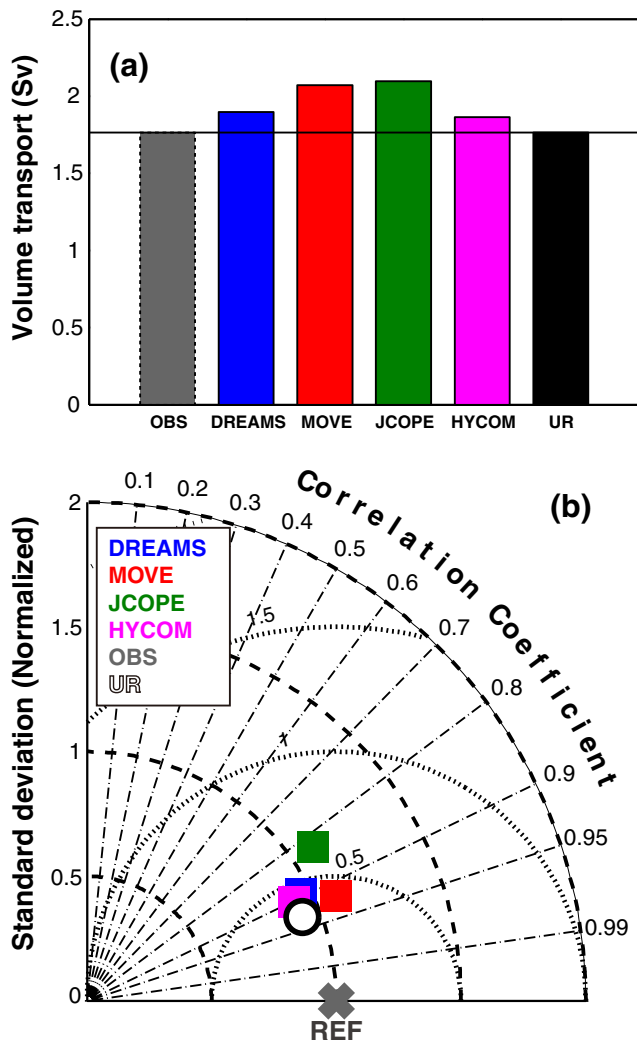
Figure 9 shows the monthly averaged transports of the observed data, the four reanalyses from the ocean models, and the UR estimate from 2003 to 2007 for each strait. The MME estimate considering a connected system tends to catch up with the observed data, such as was shown in the RR result (Fig. 6). However, a difference exists between the previous and present MME estimations despite use of the same method. The significant difference is that the transport through the TG Strait of the MME estimation with the UR is larger than that of the RR by 0.16 Sv. In addition, the volume transports through the other straits of the MME results are also changed. These changes indicate that the UR considers not only the volume transport corresponding to the strait but also the transports of other straits, improving the balance between incoming and outgoing transport compared with RR.

Figure 10a depicts the 5-year mean volume transport from 2003 to 2007 through the major straits in the EJS system. The 5-year mean transport of the MME result is remarkably similar to that of the observed data, although all of the assimilated model results are overestimated. The JCOPE model has the largest volume transport among the models, and the HYCOM model is fairly similar to the observation.

The correlation coefficient ( $R$ ), normalized STD ( $\hat{\sigma}$ ), and RMSD ( $E$ ) between the observation and the MME result and the estimations from the models are represented by the Taylor diagram in Fig. 10b. The MME point is much closer to the reference than any individual model point. The correlation coefficient of the MME result is 0.97, whereas the correlation coefficients between observation and the individual models (DREAMS, MOVE, JCOPE, and HYCOM) are 0.90, 0.92, 0.82, and 0.90, respectively. Thus, the UR result also has the highest correlation coefficient with the observed data. The normalized STD is slightly smaller than unity, and the DREAMS and HYCOM points also show normalized STD less than unity. According to the three statistics in Fig. 10, the consolidation model most resembles the observation. The resulting equation for the MME is

$$y_{\text{Transport}} \approx -0.053 + 0.503 x_{\text{DREAMS}} + 0.503 x_{\text{MOVE}} - 0.352 x_{\text{JCOPE}} + 0.301 x_{\text{HYCOM}} \quad (11)$$

where the intercept coefficient shown as the first term is  $-0.053$ . It means that all models have a bias as the amount of the intercept coefficient. Although not shown in Eq. (11), the residual is almost zero. This error term also contributes to the bias of MME model. The mean squared error (MSE) as variance of the errors is  $0.08 \text{ Sv}^2$ . The MSE suggests that the MME models could estimate the volume transport of the EJS compared with single models (Fig. 7d–f). If the regression



**Fig. 10** (a) Five-year mean of volume transport. The line is the same as the observation value. (b) Statistical analysis using the Taylor diagram for the volume transports considered in the connected system

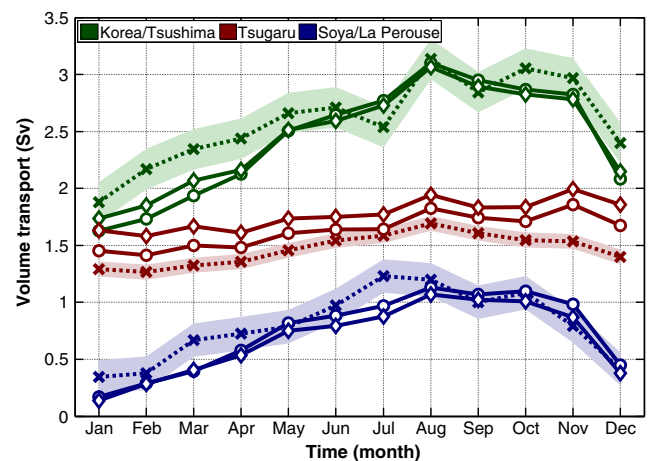
coefficients of the four models are arranged in the sequence of weights, we obtain DREAMS and MOVE, HYCOM, and JCOPE. The sequence of the regression coefficient is exposed by the three normalized statistics and the five mean transports of the four models (Fig. 10). The DREAMS and MOVE models, which have the large regression coefficient, are very near unity in the normalized STD and have the highest correlation with the reference. On the other hand, the JCOPE model point is farther away from the reference in the Taylor diagram, and the 5-year mean also shows the largest difference between the line indicating the observed mean transport.

Compared with the previous regression equations (RR), which are expressed in Table 3, we find a significant improvement in the present equation (UR) as Eq. (11). First, the remarkable change, of course, is the common regression coefficients that can be applied to any strait. Second, the sum of the weights approaches unity. The total weights of the RR result were 1.268, 0.903, and 1.205 at the KT, TG, and SP Straits,

respectively, but that of the UR is 0.930. If the total weight is larger or smaller than unity, it indicates that the independent variables tend to be underestimated or overestimated from the reference. The total weights of the UR result is close to unity; thus, water mass balance is well established by the present condition. Third, the residuals approached zero. The residuals of the RR result were far from zero, 0.429, 0.283, and 0.189 for the KT, TG, and SP Straits, respectively, whereas the residuals of the UR result are almost zero. This means that the present combination model is able to estimate the observed data statistically. Moreover, the correlation coefficients between the UR and the observed data are higher than those of the RR. The correlation coefficients between the present MME estimates of 0.939, 0.871, and 0.914 are higher than those of the previous estimates of 0.893, 0.855, and 0.666 for the KT, TG, and SP Straits, respectively. Judging from these analyses, it is clear that the UR estimates for the present condition outperform the former RR estimates because the combination model with the present condition satisfies not only physical conservation but also the statistical condition.

Figure 11 shows the seasonal variations of volume transport with the MME estimates compared with observed data for the three straits to interpret the consolidation model. The KT transport of the UR estimate is smaller than the observation, except during summer. The volume transports of UR and the observed data tend to be relatively similar for the SP Strait. However, the UR estimate of TG transport is overestimated compared with the observed data.

The annual mean transports and the seasonal variability of the observed data and the consolidation estimate are quantified in Table 4. The mean transport of the obtained model for the KT Strait tends to be smaller than the measurement data by about 0.17 Sv, whereas the mean transport of the consolidation model through the TG Strait shows the opposite tendency



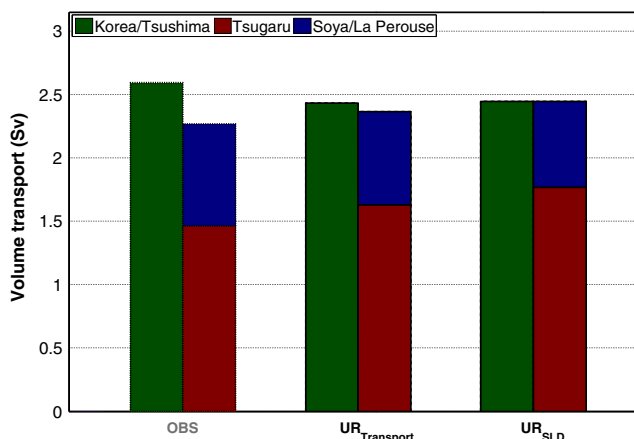
**Fig. 11** Monthly volume transport of the observed data (cross mark), consolidation estimates with transport (circle mark), and sea level difference (diamond mark) in the Korea/Tsushima (green), Tsugaru (red), and Soya/La Perouse (blue) Straits from 2003 to 2007. The light shading denotes the 90 % confidence interval of observed transports

**Table 4** Annual mean and standard deviation of observed data and multi-model ensembles for the Korea/Tsushima, Tsugaru, and Soya Straits from 2003 to 2007

Names of strait	OBS	$y_{\text{Transport}}$	$y_{\text{SLD}}$
Mean			
Korea/Tsushima	2.60	2.43	2.45
Tsugaru	1.47	1.63	1.77
Soya/La Perouse	0.80	0.74	0.68
STD			
Korea/Tsushima	0.36	0.49	0.42
Tsugaru	0.13	0.14	0.13
Soya/La Perouse	0.30	0.33	0.30

with an overestimation of about 0.16 Sv. The annual transport of the SP Strait is slightly smaller than the observed data. Compared with the RR estimates applied in the single system, they show similar trends with UR estimate for the present condition in all strait, although the difference of the MME estimates between unified and single cases for the KT Strait is relatively larger than other straits. The difference is almost 0.15 Sv, and it shows in winter–spring. On the other hand, the STDs of  $y_{\text{transport}}$  are almost identical to the observed, except for the KT Strait. In the KT Strait, the STD of the UR estimate ( $y_{\text{transport}}$ ) presents a marked contrast to that of the observed data, differing by about 0.13 Sv. These analyses indicate that the mean and seasonal variation of the observed volume transport is overestimated in the KT Strait, whereas the observed mean transport of the TG Strait is underestimated. The consolidation and observed results have almost identical mean and seasonal variation of volume transport through the SP Strait.

The inconsistency of the combination model and the observed data is shown in Fig. 12. The inconsistency of the



**Fig. 12** Differences between the transport entering through the Korea/Tsushima Strait (green) and the outgoing transport at the Tsugaru (blue) and Soya/La Perouse (red) Straits for observed data and multi-model ensemble estimations

observed data is quite large, about 0.33 Sv, but the budget inconsistency of the  $y_{\text{transport}}$  is relatively small. This demonstrates that the UR estimate is balanced with respect to the incoming transport at the KT Strait and the outgoing transport through the TG and SP Straits.

## 5 Validation of multi-model ensemble

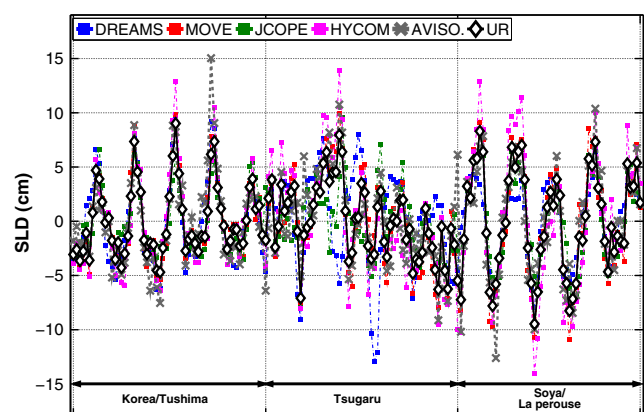
So far, the MME estimates have been obtained based on transport data. The aim of this section is to verify the UR result by using the independent variable SLD. Furthermore, we predict the seasonal variation of volume transport with the obtained regression equations from volume transport and SLD data after 3 years to validate the MME results.

### 5.1 MME with altimeter data

If another variable that is related to the original variable derives the same result, the methodology could ensure the validity of the original result. Numerous studies have shown that the variation of volume transport is closely related to the SLD between two oceans connected by shallow straits. Thus, the along-strait SLD from the satellite altimeter, which is related to the variation of transport, was considered as an independent variable.

Although the observed transport data considered a connected system have missing data for 40 months, the altimeter SLD data provide an almost continuous record (Figs. 3 and 4). The same number of ensembles of each strait may be attributed to the fair MME estimation of seasonal variation equally weighted to the three straits.

The monthly averaged SLDs of the altimeter data and four modeled results in the KT, TG, and SP Straits are shown in Fig. 13. The HYCOM reanalysis shows relatively large



**Fig. 13** Monthly sea level difference at the major straits in the East/Japan Sea. Satellite altimeter data, reanalyses from the four models, and ridge regression performance are represented by the gray cross mark, colored square mark, and black diamond mark, respectively

fluctuations, whereas the JCOPE result shows moderate change. Each reanalysis represents the seasonal cycle in the KT and SP Straits, but the seasonal variation of the TG Strait is smaller than that of the other straits.

Comparing the AVISO data with the MME conducted using SLD, the MME estimate better tracks the AVISO data as the dependent variable compared with any individual model. This evidence is shown in the Taylor diagram of Fig. 14. The RMS of the MME (in detail UR) is smaller than that of any single model result. The MME also has the highest correlation with the AVISO data (reference) compared with the four modeled points. Although the normalized STD of the MOVE model is the nearest to unity, 1.004, the MME point is also close to unity.

These normalized statistics are reflected in the regression coefficients of the regression equation. The resulting regression equation using SLD data is

$$y_{SLD} \approx 0.000 + 0.069 x_{DREAMS} + 0.464 x_{MOVE} + 0.142 x_{JCOPE} + 0.228 x_{HYCOM} \quad (12)$$

The shorter the distance between the model point and the reference in the Taylor diagram, the larger the regression coefficient gained in the regression equation. The MOVE model point has the shortest distance from the reference among the model points. Accordingly, the regression coefficient of the MOVE is the largest among the individual models. The regression coefficient of the DREAMS model is the smallest among the models because the DREAMS point is farthest from the reference.

In the MME equation with SLD, the intercept coefficient is near zero, and the sum of the regression coefficients is almost unity, about 0.903. As described in Section 4.3, this indicates

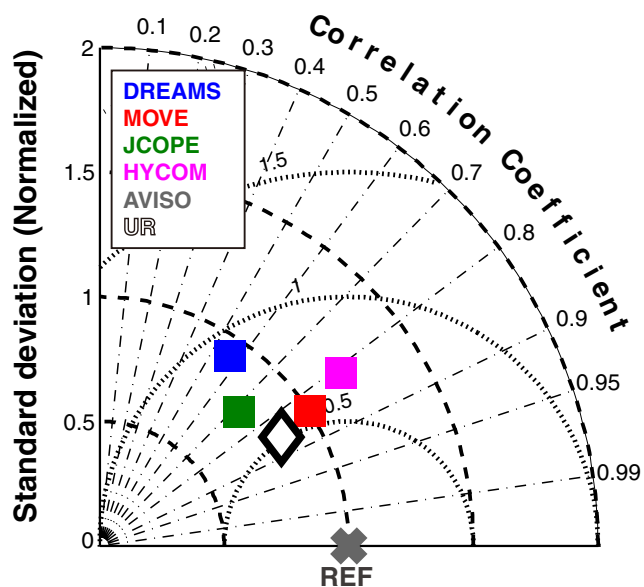


Fig. 14 Same as the lower panel of Fig. 10, but for sea level difference instead of volume transport

that the MME result satisfies not only mass conservation but also the statistical condition.

Figure 11 shows the seasonal variation of the volume transport with the observed data and the consolidation model for the equation obtained in Eq. (12) using the transport of the four assimilation models. The applied transports for the consolidated MME result tend to be smaller than the measured data in the KT Strait, except during summer, whereas those of the TG Strait are larger than the observed data.

The annual mean transport and the STD of the observed data and the estimates of the applied consolidation equations are summarized in Table 4. The mean transport of the applied MME result is smaller than the observed data in the KT Strait but larger than the observed data in the TG Strait. In the SP Strait, the mean transport of the consolidation model is smaller than the observed data, but this difference, 0.12 Sv, is relatively small.

In comparison with the STD of the observed data, the MME result has larger STD than the observed data by about 0.06 Sv in the KT Strait. For the other straits, the MME result and the observed data have almost identical STD. This indicates that the linear equation to calculate the transport of the TG Strait is able to estimate the seasonal variation of volume transport.

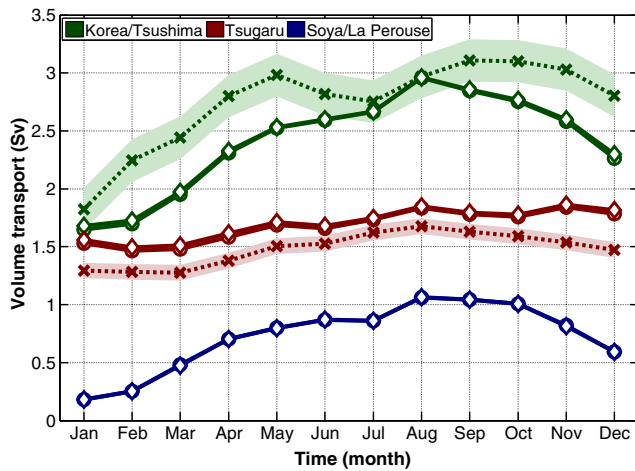
The consolidation models applied in Eqs. (11) and (12) show similar tendencies of annual mean transport and STD. The mean transports of both combination models are smaller than the observed data for the KT Strait but larger than the observed data for the TG Strait. The resemblance between the two MME results is shown by the STD. The similarity of the two consolidation models enhances the reliability of the MME estimation.

The inconsistency between incoming transport and outgoing transport is shown in Fig. 12. The inconsistency in the budget of the  $UR_{SLD}$  is smaller than that of the observed data, which indicates that the mass balance is maintained in the MME result of SLD. As shown in the MME estimation results, the inconsistency of observed data originated from the transports of the KT and TG Straits.

### 5.2 Prediction

For additional validation of the MME results, we attempted to predict the volume transport after the three years from 2008 to 2010. The prediction was carried out using transport data from the four different models with the obtained regression equations.

Figure 15 shows the seasonal variation of the volume transports through the KT, TG, and SP Straits from the observed data and the prediction. The observed data for the KT and TG Straits were available during a set period, but observed data of the SP current transport were unavailable (Fig. 2). There seem to be similar seasonal tendencies between the predictions and the observed data for the KT and TG Straits. However, the predictions are quantitatively dissimilar to the observed data. The predictions for the KT Strait tend to be smaller than the



**Fig. 15** Monthly averaged transport through the Korea/Tsushima (green), Tsugaru (red), and Soya/La Perouse (blue) Straits from 2008 to 2010. Observed volume transport is represented by the cross mark. The prediction results represented by the circle and diamond mark curves indicate the application with Eqs. (1) and (2), respectively. The shade denotes the 90 % confidence interval of observed transports

measurement data, whereas those of the TG Strait are larger than the observed data.

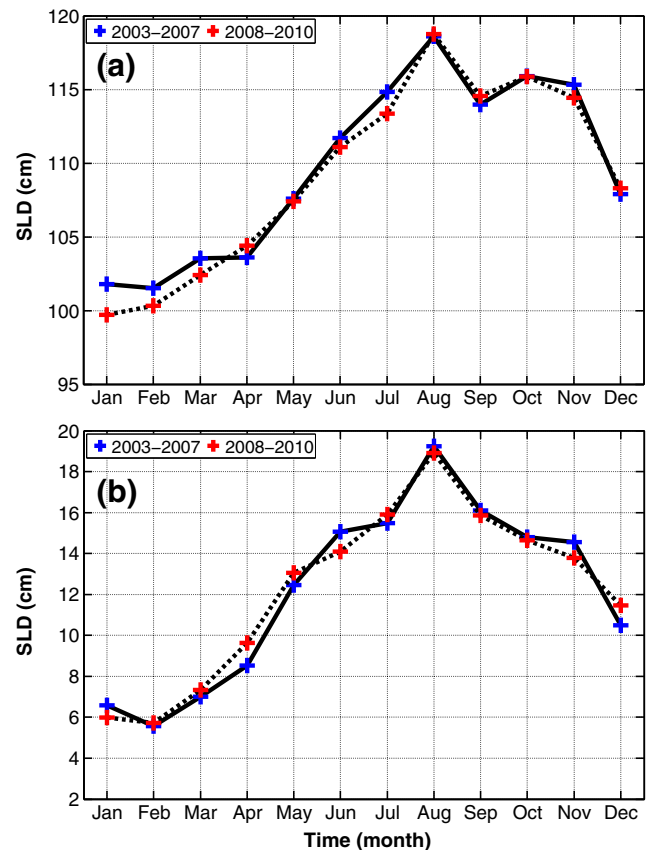
Table 5 summarizes the annual mean transports and the STDs of the observations and predictions in the three straits. The mean transports of the predictions in the KT Strait are smaller than the ADCP observation by 0.34 Sv, while those of TG Strait are larger than the observation by about 0.20 Sv. The difference between the observed data and the prediction is twofold that between the observed data and the MME estimate for the KT Strait, whereas the differences between the observed data and both the prediction and the MME estimate are generally comparable. The STDs of the volume transports also show the same tendencies as the MME estimates. The STDs of the predictions in the KT Strait show relatively large differences between the observed data compared with the other straits. Hence, the predictions are also underpinned by the MME results.

**Table 5** Annual mean and standard deviation of observed data and predictions for the Korea/Tsushima, Tsugaru, and Soya Straits from 2008 to 2010

Names of strait	OBS	$\bar{y}_{Transport}$	$\bar{y}_{SLD}$
<b>Mean</b>			
Korea/Tsushima	2.74	2.40	2.42
Tsugaru	1.48	1.68	1.70
Soya/La Perouse		0.72	0.72
<b>STD</b>			
Korea/Tsushima	0.37	0.42	0.41
Tsugaru	0.14	0.13	0.12
Soya/La Perouse		0.28	0.28

The unavailable data of the observed SP transport can be estimated by the difference in transport between the KT and TG Straits of the MME result because mass conservation is guaranteed in the predictions.

Previous studies have shown that the volume transport is related to the across-strait SLD. We compared the SLD during the set periods of the MME and prediction. The monthly SLDs across the KT and TG Straits during 2003–2007 (the MME period) and 2008–2010 (the prediction period) are shown in Fig. 16. The SLDs of the two separate periods are very similar, which suggests that the volume transports of the MME and prediction periods should also be similar. The TG observed transports in the two separate periods are almost identical quantitatively, but the observed value for the KT Strait from 2008 to 2010 is larger by 0.14 Sv compared to the 2003–2007 period. It may be that the observed transport from 2008 to 2010 is more overestimated (particularly for 2008) than in the MME period of 2003–2007. This result implies that the long-term ADCP data in the KT Strait include some error; thus, there is a need to verify the observation in not only the current period but also in previous or subsequent periods.



**Fig. 16** Comparison of the monthly sea level differences in 2003 to 2007 (red cross) and 2008 to 2010 (blue cross) across the Korea/Tsushima (a) and Tsugaru (b) Straits



## 6 Concluding remarks

The goal of this study was to estimate the transports of the KT, TG, and SP Straits for physically consistent circulation in the EJS using the MME approach. The MMEs arising from a combination of multi-models outperformed the reanalyses from the individual models by considering model uncertainties. In particular, the ridge estimator overcame the problem of multicollinearity compared with the OLS estimator.

The MME equation, which satisfies the physical and statistical conditions of these straits, was obtained using the transport data. To validate this MME estimate, the MME was carried out with SLD, which is related to transport, as the independent variable.

Comparing the two regression equations for the transport and SLD data, the estimates of MME with SLD were found to be similar to the MME results with transport data. The MOVE model was allocated the highest weight, whereas the JCOPE model had small regression coefficients in both of the consolidation models. This result demonstrates that the MOVE model can simulate the flow system of the EJS compared with the other models. The DREAMS model has strong weight in the regression equation for the transport data, even with relatively coarse horizontal grid resolution.

The MME estimate indicates that the volume transport was smaller than the measurement data at the KT Strait, but was larger than the observed data at the TG Strait. The optimal transports considering mass conservation for 2003 to 2007 were 2.43, 1.63, and 0.74 Sv for the KT, TG, and SP Straits, respectively. The MME results also suggest that STD in the KT Strait is larger than observed, whereas the estimated results are almost the same as the ones observed in the TG and SP Straits. The MME estimates for SLD data and prediction were also found to be similar to the original case with transport. These similarities enhance the reliability of the MME estimates for transport data.

Since the insufficiency of the long-term and simultaneous observational data in the three straits, it has been difficult to propose the inflow and outflow systems of the EJS until now. Even though, the ratio of the model reanalyses data, which is sufficient for the conserved transport, is different depending on the model. The outflow partitioning of four reanalyses showed that 75, 69, 82, and 73 % of the total inflow transport flows out of the EJS through the TG Strait and 25, 31, 18, and 27 % through the SP Strait in the DREAMS, MOVE, JCOPE, and HYCOM, respectively. According the present MME result, the ratio of the outflow through the TG Strait versus SP Strait is 0.68:0.32. This ratio is relatively close to Na et al. (2009) or the ratio of MOVE, which has large regression coefficient.

These results suggest the need to modify the observed transport data in the three straits to estimate physically consistent transport in the EJS system. In the case of the KT Strait, the

measured transport has been calculated from current data obtained from the vessel-mounted ADCP. When the KT current was observed on the cruise, the sampling error of time intervals or spatial points and cruising speed might have affected the accuracy of the dataset. In addition, programming errors from missing data and the calculation process need to be considered.

For the TG Strait, the similarity between the MME result and the observed data indicates that seasonal variation in volume transport can be simulated with the along-strait SLD. The equation used to estimate the observed transport of the TG Strait was  $Q=0.0271\Delta\eta+0.9333$  (Nishida et al. 2003). The linear equation is based on the ADCP data observed from 1993 to 1999. Nevertheless, this equation was able to estimate the seasonal variation in the volume transport of the TG Strait in other periods, whereas the mean transport of MME estimate has been underestimated compared with any single model results, at least in 2003 to 2010. This indicates that the estimation equation is necessary to consider the interannual variability, that is, the need exists to increase the term of the  $y$ -intercept in this equation.

Although these MME results are consistent with mass conservation in the EJS circulation, we do not explain why the TG transports from all model reanalyses have a large gap within the observed data. The dynamic mechanism that overestimated the volume transport of the TG strait will be investigated in a future study. Overall, this paper suggests that the volume transports of the KT, TG, and SP Straits are physically consistent.

**Acknowledgments** This study was supported by the MEXT Special Funds for Education and Research (Joint Usage/Research Center).

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

- Bang I, Choi J-K, Kantha L, Horton C, Clifford M, Suk M-S, Chang K-I, Nam SY, Lie H-J (1996) A hindcast experiment in the East Sea (Sea of Japan). *La mer* 34:108–130
- Chang K-I, Teague WJ, Lyu SJ, Perkins HT, Lee D-K, Watts DR, Kim Y-B, Mitchell DA, Lee CM, Kim K (2004) Circulation and currents in the southwestern East/Japan Sea: overview and review. *Prog Oceanogr* 61:105–156. doi:10.1016/j.pocean.2004.06.005
- Chu PC, Lan J, Fan C (2001a) Japan sea thermohaline structure and circulation. Part II: a variational P-vector method. *J Phys Oceanogr* 31:2886–2902. doi:10.1175/1520-0485(2001)031<2886:JSTSAC>2.0.CO;2
- Chu PC, Lu SH, Fan CW, Kim CS (2001a) Modeling of Japan Sea circulation and thermohaline structure. *Advances in coastal modeling*, V. C. Lakhan, Ed., Elsevier Science (in press)

- Csanady GT (1982) Circulation in the coastal ocean. Reidel, Dordrecht, p 279
- Fukamachi Y, Ohshima KI, Ebuchi N, Bando T, Kazuya O, Minoru S (2010) Volume transport in the Soya Strait during 2006–2008. *J Oceanogr* 66(5):685–696. doi:10.1007/s10872-010-0056-2
- Fukudome KI, Yoon JH, Ostrovskii A, Takikawa T (2010) Seasonal volume transport variation in the Tsushima Warm Current through the Tsushima Straits from 10 years of ADCP observations. *J Oceanogr* 66:539–551. doi:10.1007/s10872-010-0045-5
- Garrett C, Petrie B (1981) Dynamical aspects of the flow through the Strait of Belle Isle. *J Phys Oceanogr* 11:376–393. doi:10.1175/1520-0485(1981)011<0376:DAOTFT>2.0.CO;2
- Hagedorn R, Doblas-Reyes FJ, Palmer TN (2005) The rationale behind the success of multi-model ensembles in seasonal forecasting—I. Basic concept. *Tellus A* 57(3):219–233. doi:10.1111/j.1600-0870.2005.00103.x
- Hata K (1973) Variations in hydrographic conditions in the seas adjacent to the Tsugaru Strait. *J Meteorol Res* 25:467–473 (in Japanese)
- Hoerl AE, Kennard RW (1970) Ridge regression: biased estimation for nonorthogonal problems. *Technometrics* 12(1):55–67
- Hoerl AE, Kannard RW, Baldwin KF (1975) Ridge regression: some simulations. *Commun Stat* 4(2):105–123
- Isobe A (1994) On the Tsushima warm current in the Tsushima strait. *Kaiyo Monthly* 26:802–809 (in Japanese)
- Ito T, Togawa O, Ohnishi M, Isoda Y, Nakayama T, Shima S, Kuroda H, Iwahashi M, Sato C (2003) Variation of velocity and volume transport of the Tsugaru Warm Current in the winter of 1999–2000. *Geophys Res Lett* 30(13):1678–1681. doi:10.1029/2003GL017522
- Krishnamurti TN, Kishtawal CM, Zhang Z, LaRow T, Bachiochi D, Williford E, Gadgil S, Surendran S (1999a) Multi-model superensemble forecasts for weather and seasonal climate FSU (Florida State Univ.) Report 99–8, July 1999
- Krishnamurti TN, Kishtawal CM, LaRow T, Bachiochi D, Zhang Z, Williford E, Gadgil S, Surendran S (1999b) Improved weather and seasonal climate forecast from multimodel superensemble. *Science* 285(5433):1548–1550
- Krishnamurti TN, Kishtawal CM, Zhang Z, LaRow T, Bachiochi D, Williford E, Gadgil S, Gadgil S, Surendran S (2000) Multimodel ensemble forecasts for weather and seasonal climate. *J Clim* 13:4196–4216. doi:10.1175/1520-0442(2000)013<4196:MEFFWA>2.0.CO;2
- Lyu SJ, Kim K (2003) Absolute transport from the sea level difference across the Korea Strait. *Geophys. Res Lett* 30(6):18-1–18-4. doi:10.1029/2002GL016233
- Matsuyama M, Wadaka M, Abe T, Aota M, Koike Y (2006) Current structure and volume transport of the Soya Warm Current in summer. *J Oceanogr* 62:197–205
- Na H, Isoda Y, Kim K, Kim YH, Lyu SJ (2009) Recent observations in the straits of the East/Japan Sea: a review of hydrography, currents and volume transports. *J Mar Syst* 78(2):200–205. doi:10.1016/j.jmarsys.2009.02.018
- Nishida Y, Kanomata I, Tanaka I, Sato S, Takahashi S, Matsubara H (2003) Seasonal and interannual variations of the volume transport through the Tsugaru Strait. *Umi no Kenkyu* 12:487–499 (in Japanese with English abstract)
- Ohshima KI (1994) The flow system in the Japan Sea caused by a sea level difference through shallow straits. *J Geophys Res* 99(C5):9925–9940. doi:10.1029/94JC00170
- Pearson K (1908) On the generalized probable error in multiple normal correlation. *Biometrika* 6:59–68
- Peña M, Van den Dool H (2008) Consolidation of multimodel forecasts by ridge regression: application to Pacific sea surface temperature. *J Clim* 21:6521–6538. doi:10.1175/2008JCLI2226.1
- Peng P, Kumar A, van den Dool H, Barnston AG (2002) An analysis of multimodel ensemble predictions for seasonal climate anomalies. *J Geophys Res* 107(D23): ACL 18-1–ACL 18-12. doi:10.1029/2002JD002712
- Seung YH, Han SY, Lim EP (2012) Seasonal variation of volume transport through the straits of the East/Japan Sea viewed from the island rule. *Ocean Polar Res* 34(4):403–411. doi:10.4217/OPR.2012.34.403
- Takikawa T, Yoon JH, Cho KD (2005) The Tsushima warm current through Tsushima straits estimated from ferryboat ADCP data. *J Phys Oceanogr* 35:1154–1168. doi:10.1175/JPO2742.1
- Taylor KE (2000) Summarizing multiple aspects of model performance in a single diagram. PCMDI Report No 65., Program for Climate Model Diagnosis and Intercomparison, Lawrence Livermore National Laboratory, University of California, Livermore CA, 24 pp
- Teague WJ, Jacobs GA, Perkins HT, Book JW, Chang K-I, Suk M-S (2002) Low-frequency current observations in the Korea/Tsushima Strait. *J Phys Oceanogr* 32(6):1621–1641. doi:10.1175/1520-0485(2002)032<1621:LFCOIT>2.0.CO;2
- Toba Y, Tomizawa K, Kurasawa Y, Hanawa K (1982) Seasonal and year-to-year variability of the Tsushima-Tsugaru warm current system with its possible cause. *La Mer* 20:41–51
- Toulany B, Garrett CJR (1984) Geostrophic control of fluctuating flow through straits. *J Phys Oceanogr* 14:649–655
- Yanagi T (2002) Water, salt, phosphorus and nitrogen budgets of the Japan Sea. *J Oceanogr* 58(6):797–804. doi:10.1023/A:1022815027968