3rd Japanese DA workshop, Jan 18th, 2013, ISM, Tachikawa

Pathways of the North Pacific Intermediate Water identified through the tangent linear and adjoint models of an OGCM

Presenter: Yosuke Fujii (JMA/MRI)

Coauthor: Toshiya Nakano[*1], Norihisa Usui[*2], Satoshi Matsumoto[*1], Hiroyuki Tsujino[*2], and Masafumi Kamachi[*2]

[*1]JMA, [*2]JMA/MRI

Fujii et al. 2013, JGR, accepted

★ Outline

Part I: Theoretical Introduction

1 Derivation of an adjoint model
2 Backward tracing using an adjoint model
3 Joint likelihood and Probability Density (PD)

Part II: Practical Application to NPIW.

4 Purpose of tracing NPIW
5 Results (Pathway of NPIW from the subarctic region to the south of Japan)
6 Discussions (Is the pathway realistic?)

7 Concluding Remarks
(Benefits from adjoint models)

Part I: Theoretical Introduction

1 Derivation of an adjoint model

★ Derivation of adjoint model (I)

Consider a state vector, $\mathbf{x}$, containing n *variables (all variables have 3D).*

e.g. $\mathbf{x} = (x_{1,t}, x_{2,t}, x_{3,t}, x_{4,t}) = (T, S, u, v)$

The forecast model of $\mathbf{x}$: $\frac{\partial \mathbf{x}}{\partial t} = M(\mathbf{x})$

Consider a cost function, $J[\mathbf{x}(\tau)] = J[x_1(\tau), x_2(\tau), \cdots, x_n(\tau)]$

(For simplicity, we assume that J directly depends on the state vector at time τ alone.)

The 1st-order variation of J is

$$\delta J = \iiint_V \left(\frac{\partial J}{\partial x_1(\tau)} \delta x_1(\tau) + \frac{\partial J}{\partial x_2(\tau)} \delta x_2(\tau) \cdots + \frac{\partial J}{\partial x_n(\tau)} \delta x_n(\tau) \right) dV = \left\langle \frac{\partial J}{\partial \mathbf{x}(\tau)}, \delta \mathbf{x}(\tau) \right\rangle = \langle \boldsymbol{\lambda}(\tau), \delta \mathbf{x}(\tau) \rangle$$

Definition of the inner product: $\langle \mathbf{y}, \mathbf{x} \rangle = \iiint_V (y_1 x_1 + y_2 x_2 \cdots + y_n x_n) dV$

Forward variables: $\delta \mathbf{x}(t)$, Adjoint variables: $\boldsymbol{\lambda}(t) = \frac{\partial J}{\partial \mathbf{x}(t)} = \left(\frac{\partial J}{\partial x_1(t)}, \frac{\partial J}{\partial x_2(t)}, \cdots, \frac{\partial J}{\partial x_n(t)} \right)$

☆ Derivation of adjoint model (II)

If $\mathbf{x}$ is specified at a certain time t, the cost function J is also determined.

Therefore $$\delta J = \left\langle \frac{\partial J}{\partial \mathbf{x}(\tau)}, \delta\mathbf{x}(\tau) \right\rangle = \left\langle \frac{\partial J}{\partial \mathbf{x}(t)}, \delta\mathbf{x}(t) \right\rangle = \langle \boldsymbol{\lambda}(t), \delta\mathbf{x}(t) \rangle = \mathrm{const.}$$

This leads to

$$\langle \boldsymbol{\lambda}(t+\delta t), \delta\mathbf{x}(t+\delta t) \rangle = \langle \boldsymbol{\lambda}(t), \delta\mathbf{x}(t) \rangle$$

$$\Downarrow$$

$$\frac{\langle \boldsymbol{\lambda}(t+\delta t), \delta\mathbf{x}(t+\delta t) \rangle - \langle \boldsymbol{\lambda}(t+\delta t), \delta\mathbf{x}(t) \rangle}{\delta t} = \frac{\langle \boldsymbol{\lambda}(t), \delta\mathbf{x}(t) \rangle - \langle \boldsymbol{\lambda}(t+\delta t), \delta\mathbf{x}(t) \rangle}{\delta t}$$

$$\Downarrow$$

$$\left\langle \boldsymbol{\lambda}(t+\delta t), \frac{\delta\mathbf{x}(t+\delta t) - \delta\mathbf{x}(t)}{\delta t} \right\rangle = \left\langle \frac{\boldsymbol{\lambda}(t) - \boldsymbol{\lambda}(t+\delta t)}{\delta t}, \delta\mathbf{x}(t) \right\rangle$$

$$\Downarrow$$

$$\left\langle \boldsymbol{\lambda}(t), \frac{\partial \delta\mathbf{x}}{\partial t} \right\rangle = \left\langle \frac{\partial \boldsymbol{\lambda}}{\partial(-t)}, \delta\mathbf{x}(t) \right\rangle$$

★ Derivation of adjoint model (III)

The forecast model leads to

$$\frac{\partial(\mathbf{x}+\delta\mathbf{x})}{\partial t} = M(\mathbf{x}+\delta\mathbf{x}) = M(\mathbf{x}) + \mathbf{M}\delta\mathbf{x} \quad \Rightarrow \quad \frac{\partial \delta\mathbf{x}}{\partial t} = \mathbf{M}\delta\mathbf{x}$$

(Tangent Linear model)

$$\left\langle \boldsymbol{\lambda}, \frac{\partial \delta\mathbf{x}}{\partial t} \right\rangle = \left\langle \boldsymbol{\lambda}, \mathbf{M}\delta\mathbf{x} \right\rangle = \left\langle \mathbf{M}^{*}\boldsymbol{\lambda}, \delta\mathbf{x} \right\rangle = \left\langle \frac{\partial \boldsymbol{\lambda}}{\partial(-t)}, \delta\mathbf{x} \right\rangle$$

Use the definition of adjoint operators

Therefore, $\frac{\partial \boldsymbol{\lambda}}{\partial(-t)} = \mathbf{M}^{*}\boldsymbol{\lambda}$ (Adjoint model)

Backward time evolution of adjoint variables $\boldsymbol{\lambda}(t) = \partial J / \partial \mathbf{x}(t)$ can be calculated by the adjoint model.

In a simple 4DVAR system,

$$J = [\mathbf{x}(\tau) - \mathbf{x}_{obs}(\tau)]^{T}\mathbf{R}^{-1}[\mathbf{x}(\tau) - \mathbf{x}_{obs}(\tau)]/2$$

$$\boldsymbol{\lambda}(\tau) = \partial J / \partial \mathbf{x}(\tau) = \mathbf{R}^{-1}[\mathbf{x}(\tau) - \mathbf{x}_{obs}(\tau)]$$

→ Information of errors or data misfit

Thus, the adjoint model propagates the information of errors or data misfit backward in time.

☆ Conservation of the forward-adjoint inner product

In the derivation, we use

$$\delta J = \left\langle \frac{\partial J}{\partial \mathbf{x}(\tau)}, \delta \mathbf{x}(\tau) \right\rangle = \left\langle \frac{\partial J}{\partial \mathbf{x}(t)}, \delta \mathbf{x}(t) \right\rangle = \left\langle \boldsymbol{\lambda}(t), \delta \mathbf{x}(t) \right\rangle = \text{const.}$$

This means that

The inner product of forward and adjoint variables are conserved.

- forward: Evolved by tangent linear model
- adjoint: Evolved by adjoint model

・This fact holds for any set of tangent linear and adjoint models.

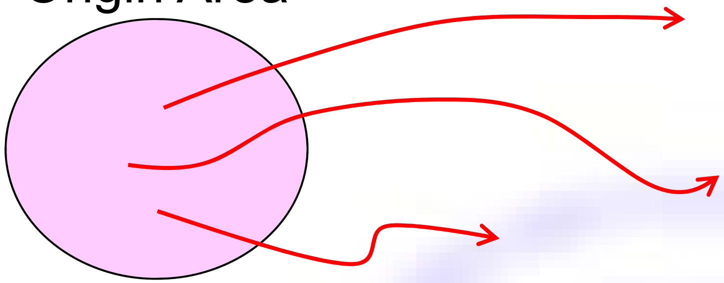
・If the cost function is defined properly, the conservation property has a physical meaning.

・In this study, we use this property for the tracing of an ocean water mass.

2 Backward tracing using an adjoint model

★ Forward and Backward Conditions

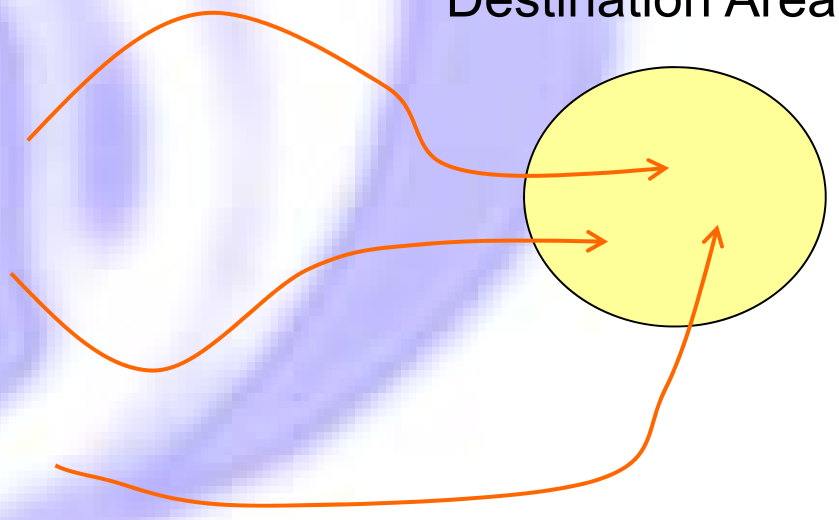
Origin Area



Forward Condition

The fluid particle is observed in the origin area at the initial time.

Destination Area



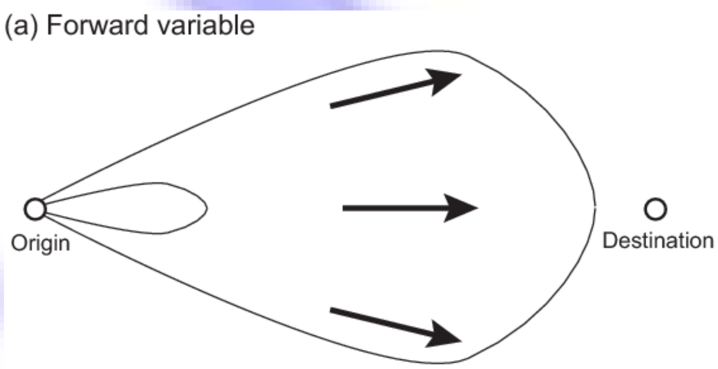
Backward Condition

The fluid particle is observed in the destination area at the final time.

★ Forward Tracing with an advection-diffusion model

・Linear advection-diffusion model for a incompressible fluid

$$\frac{\partial S}{\partial t} = -\mathbf{u} \bullet \nabla S + \nabla \bullet (\boldsymbol{\kappa} \nabla S)$$



・Now, we try to trace fluid particles that satisfy the forward condition.

If we set $S=1/V$ in the origin area
$S=0$ out of the area
at the initial time.
V: Volume of the origin area

→ Probability density (PD) of the particles that satisfy the forward condition ($P_F(\mathbf{x})$).

If we set $S=1$ in the origin area
$S=0$ out of the area

→ Fraction of the particles that satisfy the forward condition.
(i.e., Forward likelihood, $L_F(\mathbf{x})$)

Tangent linear model is written in the same form.
⟶ $P_F(\mathbf{x})$ and $L_F(\mathbf{x})$ can be considered as a forward variable.

★ Derivation of the adjoint operator

$$M\delta S = \left(\frac{\partial \delta S}{\partial t} = \right) - \mathbf{u} \bullet \nabla \delta S + \nabla \bullet \left(\boldsymbol{\kappa} \nabla \delta S\right) \quad \Rightarrow \quad \left\langle \lambda, \mathrm{M}\,\delta S \right\rangle = \left\langle \mathrm{M}^* \lambda, \delta S \right\rangle$$

$$\iiint_V \lambda \mathbf{u} \bullet \nabla \delta S + \lambda \nabla \bullet \left(\boldsymbol{\kappa} \nabla \delta S\right) dV$$

$$= \cancel{\iint [\lambda \mathbf{u} \delta S]_{x=-\infty}^{x=\infty} dydz} + \cancel{\iint [\lambda \mathbf{u} \delta S]_{y=-\infty}^{y=\infty} dzdx} + \cancel{\iint [\lambda \mathbf{u} \delta S]_{z=-\infty}^{z=\infty} dxdy} - \iiint_v \nabla \bullet (\lambda \mathbf{u}) \delta S dV$$

$$= \iiint_v (\nabla \lambda \bullet \mathbf{u}) \delta S dV + \cancel{\iiint_v (\lambda \nabla \bullet \mathbf{u}) \delta S dV} = \iiint_v \delta S (\mathbf{u} \bullet \nabla \lambda) dV$$

$$\iiint_V \lambda \nabla \bullet \left(\boldsymbol{\kappa} \nabla \delta S\right) dV = -\iiint_V \nabla \lambda \bullet \left(\boldsymbol{\kappa} \nabla \delta S\right) dV$$

$$= -\iiint_V (\boldsymbol{\kappa}^T \nabla \lambda) \bullet \nabla \delta S dV = \iiint_V \delta S \nabla \bullet (\boldsymbol{\kappa}^T \nabla \lambda) dV$$

$$\left\langle \lambda, \mathrm{M}\,\delta S \right\rangle = \iiint_V [-\lambda \mathbf{u} \bullet \nabla \delta S + \lambda \nabla \bullet \left(\boldsymbol{\kappa} \nabla \delta S\right)] dV$$

$$= \iiint_V [\delta S \mathbf{u} \bullet \nabla \lambda + \delta S \nabla \bullet \left(\boldsymbol{\kappa}^T \nabla \lambda\right)] dV = \left\langle \mathrm{M}^* \lambda, \delta S \right\rangle$$

$$\Rightarrow \quad \frac{\partial \lambda}{\partial (-t)} = \mathrm{M}^* \lambda = \mathbf{u} \bullet \nabla \lambda + \nabla \bullet \left(\boldsymbol{\kappa}^T \nabla \lambda\right)$$

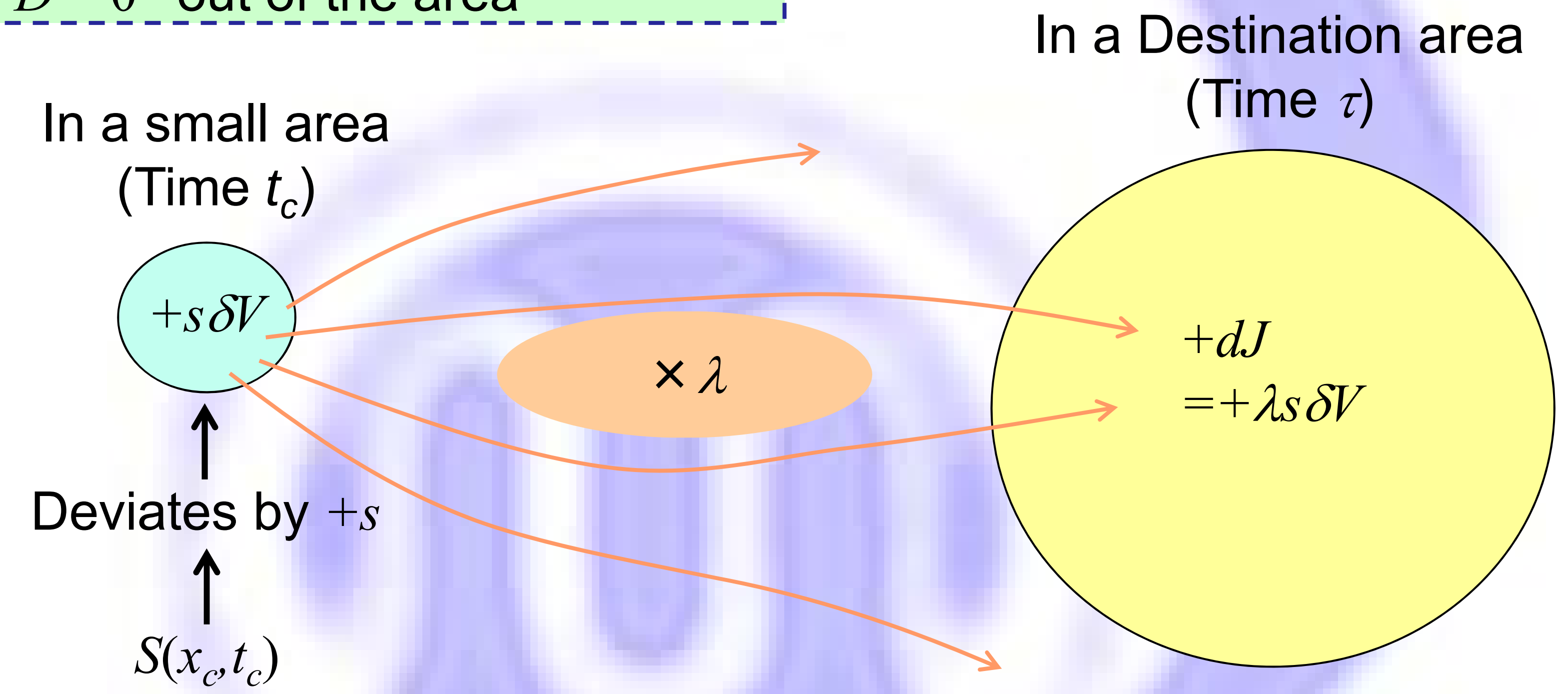
★ Backward Likelihood

Define $J = \iiint_v D S(\tau) dV = \langle D, S(\tau) \rangle$

$D = 1$ in the destination area

$D = 0$ out of the area

$\Rightarrow \quad \delta J = \langle \lambda(t), \delta S(t) \rangle$



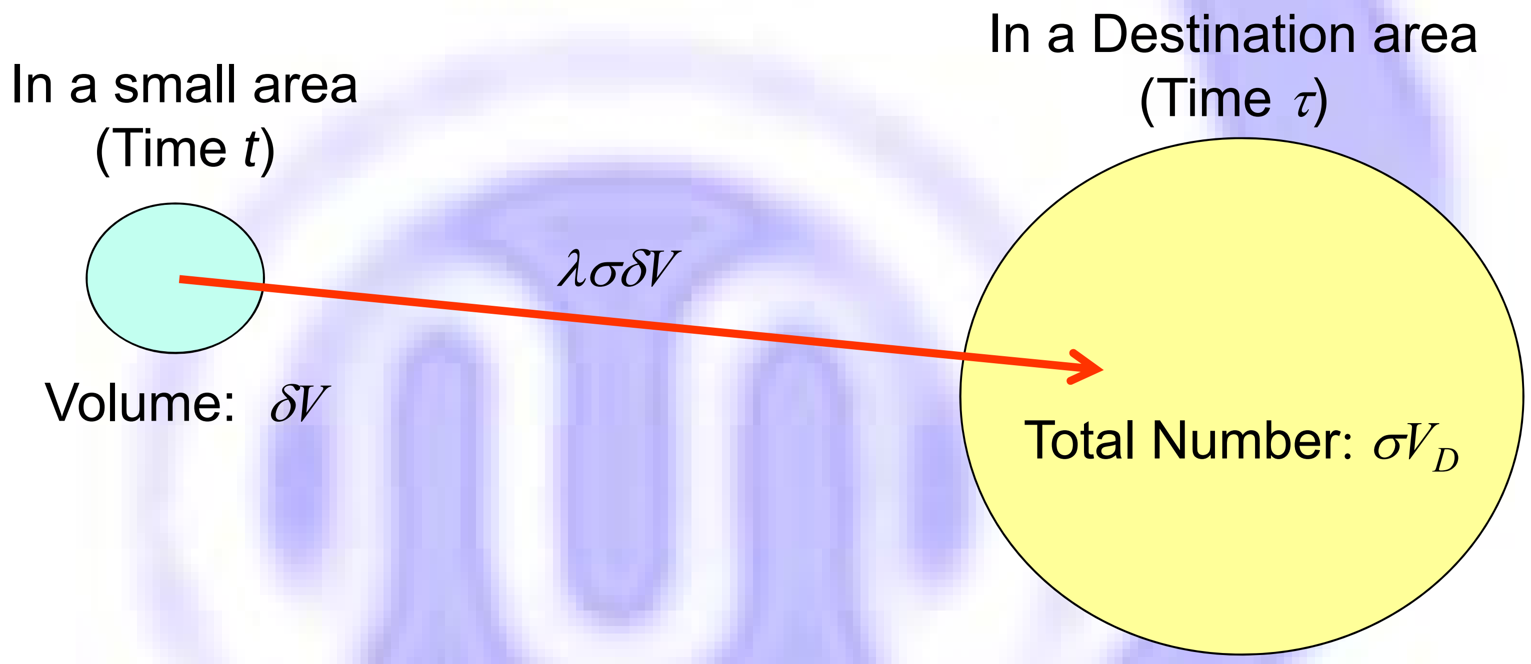
The adjoint variable $\lambda(x,t)$ can be considered as the fraction of the fluid particles that satisfy the backward condition.

We refer to this fraction as the backward likelihood, $L_B(\mathbf{x})$.

★ Backward Probability Density

Consider the case where $\delta S(t_c)=\sigma$ (σ is constant in space).

⇒ $\delta S(t)=\sigma$ for any t. ⇒ $\delta J = \langle \lambda(\tau), \delta S(\tau) \rangle = \langle D, \sigma \rangle = \sigma V_D$



PD of particles that satisfy the backward condition is represented by

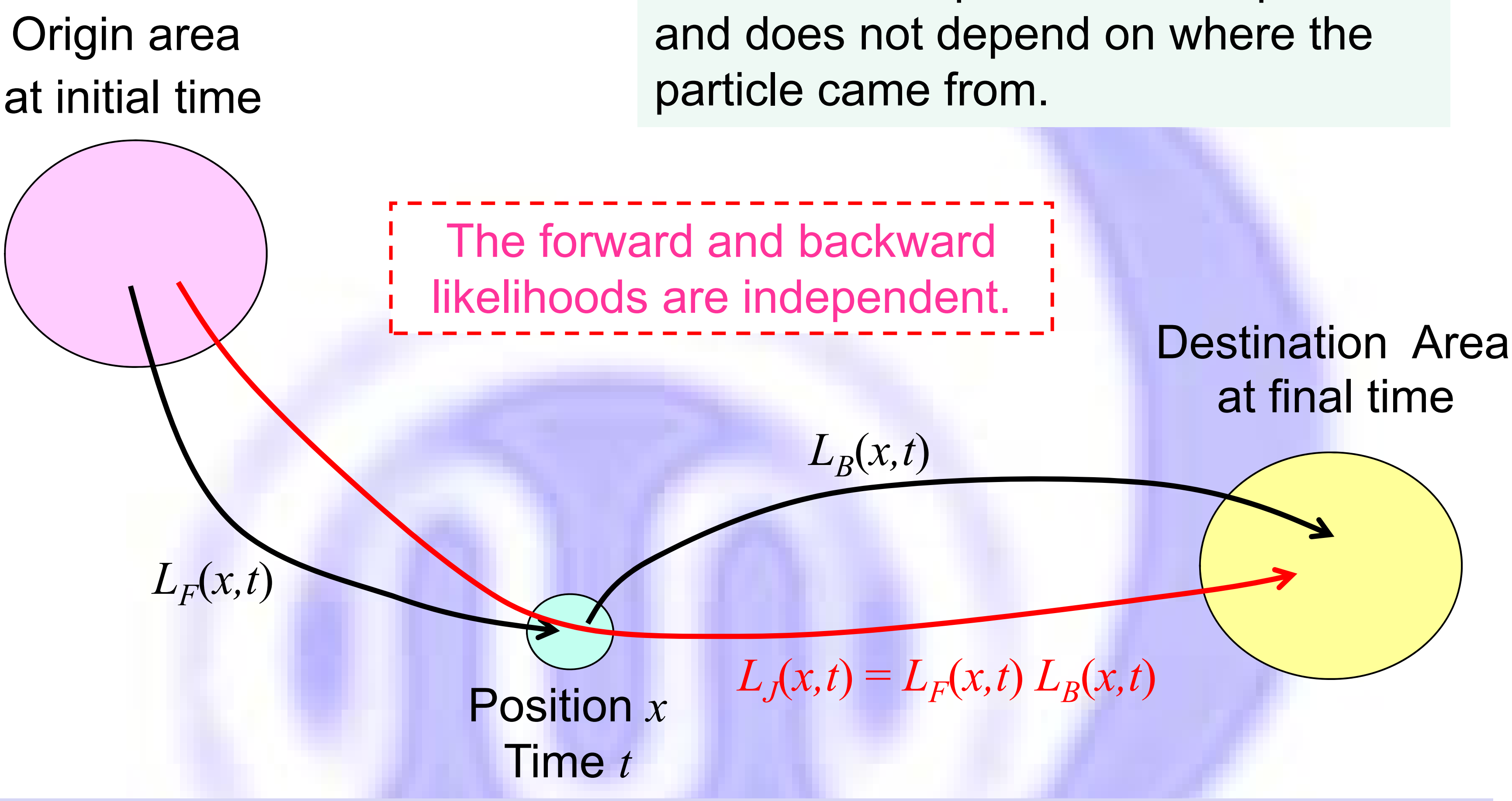
$$P_B(x,t) = \lambda(x,t)\sigma\delta V / (\sigma V_D \bullet \delta V) = \lambda(x,t)/V_D$$

We refer to this PD as the backward Probability Density.

3 Joint likelihood and Probability Density (PD)

★ Joint Likelihood

Where a particle will go depends only on the current position of the particle and does not depend on where the particle came from.



$L_J(x,t)$: Joint Likelihood (i.e., the fraction of particles that satisfy both the forward and backward condition)

The joint likelihood is expressed by the product of forward and adjoint variables.

★ Joint Provability Density (PD)

Bayes Theorem leads to $$P_J(x) = p(x \mid F, B) = \frac{p(x \mid F)\, p(B \mid x, F)}{p(B \mid F)}$$

Here, $p(x \mid F) = P_F(x)$, $p(B \mid x, F) = p(B \mid x) = L_B(x)$

and

$$p(B \mid F) = \iiint_V p(x \mid F)\, p(B \mid x, F)\, dV = \iiint_V P_F(x) L_B(x)\, dV = \langle P_F(x), L_B(x) \rangle = N_J$$

N_J: Number of particles that satisfy both forward and backward conditions.

A constant ⟷ Forward-adjoint Inner product is conserved.

Finally, $P_J(x) = P_F(x) L_B(x) / N_J$

Joint PD is also written as a product of forward and adjoint variables.

☆ Time Evolution of the joint likelihood and PD

Forward Model: $\dfrac{\partial S}{\partial t} = -\mathbf{u} \bullet \nabla S + \nabla \bullet (\boldsymbol{\kappa} \nabla S)$

Adjoint Model: $\dfrac{\partial \lambda}{\partial(-t)} = \mathbf{u} \bullet \nabla \lambda + \nabla \bullet (\boldsymbol{\kappa}^T \nabla \lambda)$

Consistent with the conservation property of forward-adjoint inner product.

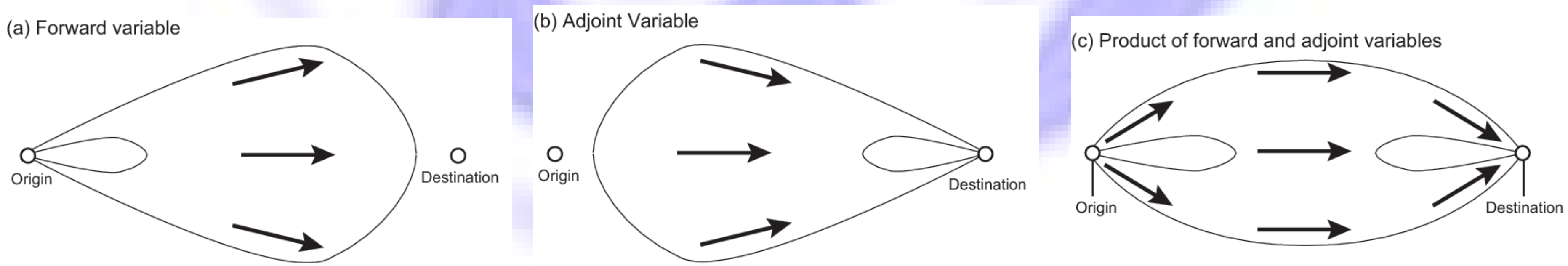
Gauss's Theorem leads to that Integration over the model domain is conserved.

Joint Probability Density

$$\frac{\partial}{\partial t}(S\lambda) = -\mathbf{u} \bullet \nabla (S\lambda) + \lambda \nabla \bullet (\boldsymbol{\kappa} \nabla S) - S \nabla \bullet (\boldsymbol{\kappa}^T \nabla \lambda) = \nabla \bullet \left(-S\lambda \mathbf{u} + \lambda \boldsymbol{\kappa} \nabla S - S \boldsymbol{\kappa}^T \nabla \lambda\right)$$

Advection term: same as in the forward model

Diffusion term: Near the initial time → The forward diffusion is dominant.
Near the final time → The adjoint diffusion is dominant.



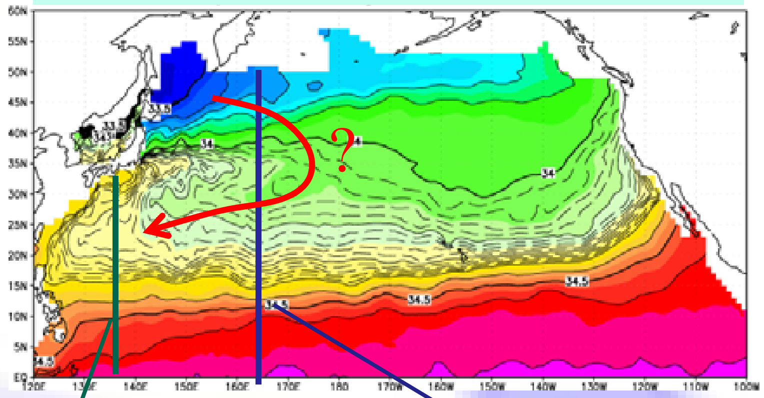
Part II: Practical Application to NPIW

4 Purpose of tracing NPIW

☆ What is NPIW?

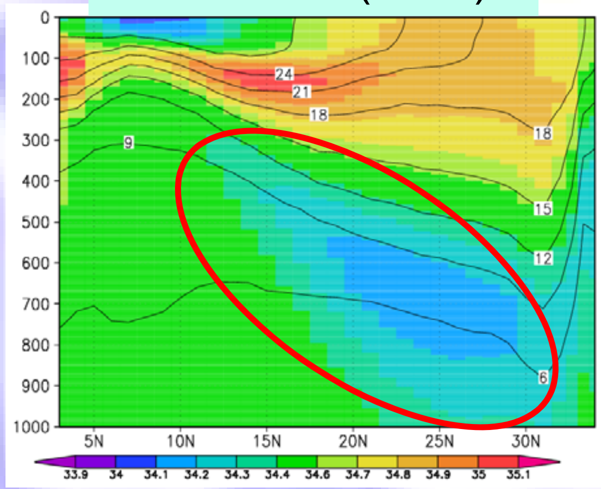
NPIW:
North Pacific Intermediate Water

On $26.8\sigma_\theta$ surface (Hydrobase)

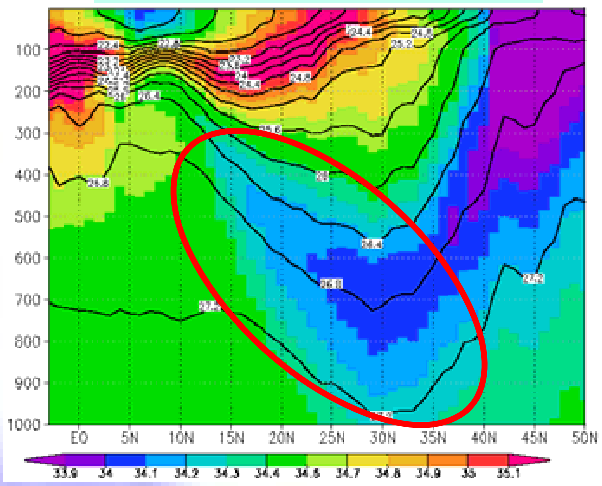


・NPIW distributes around $26.8\sigma_\theta$ surface in the North Pacific subtropical region.

137E Clim. (Obs.)



165E Clim. (Obs.)



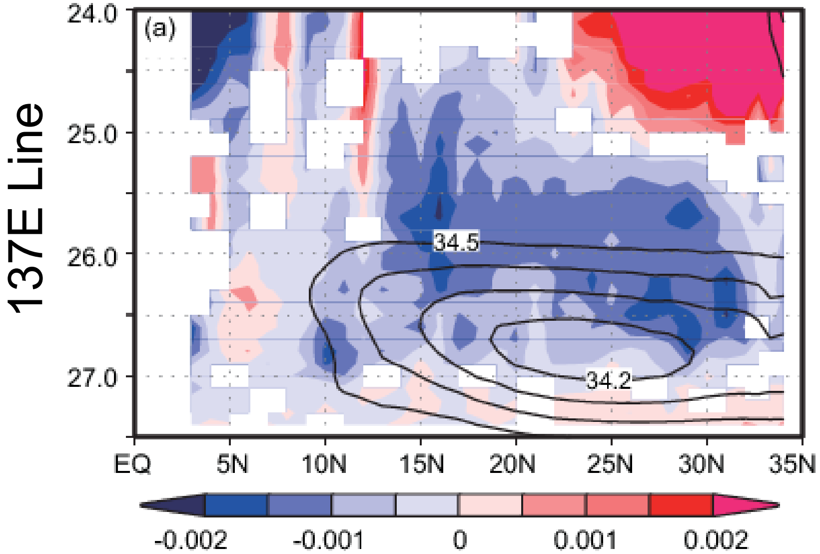
・NPIW is generated in the mixed water region and the Kuroshio Extension.

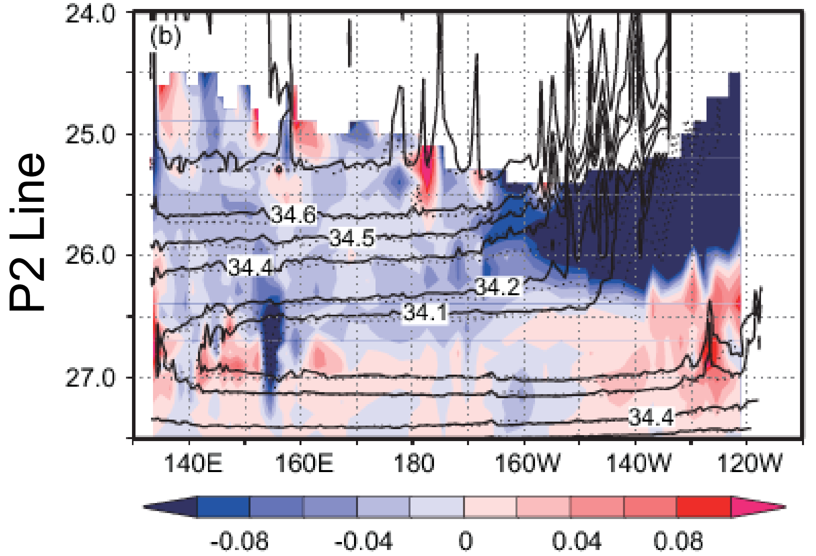
・Low salinity in NPIW originates from the near surface low salinity water in the subarctic region (Okhotsk and Bering Seas?).

・The pathway of NPIW that travels from MWR to the western part of the subtropical gyre is still controversial (e.g. Kouketsu et al. 2007, 2008).

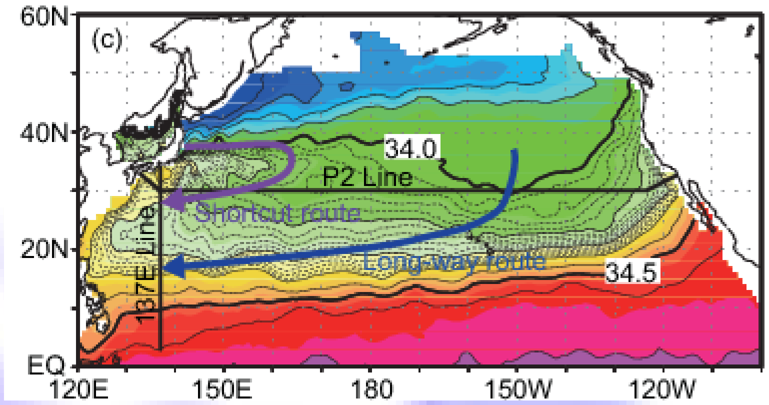
★ Freshening Trend in NPIW (from Nakano et al. 2007)

Freshening Trend in vertical sections





Salinity on $26.8\sigma_\theta$

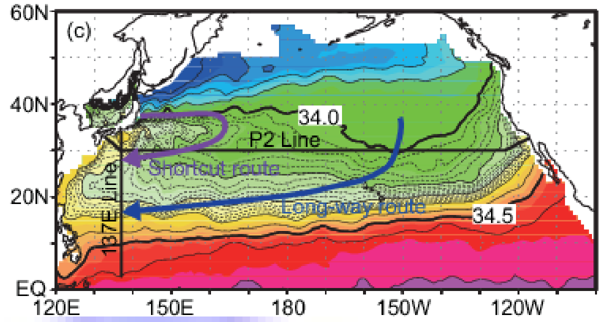


・A decadal-scale freshening trend is found in the upper part of NPIW in the 137E section.

・Kouketsu et al. (2007) also indicated a similar freshening trend of NPIW.

・Nakano et al. (2007) speculated existence of the shortcut and long-way routes of NPIW.

★ Purpose and Method

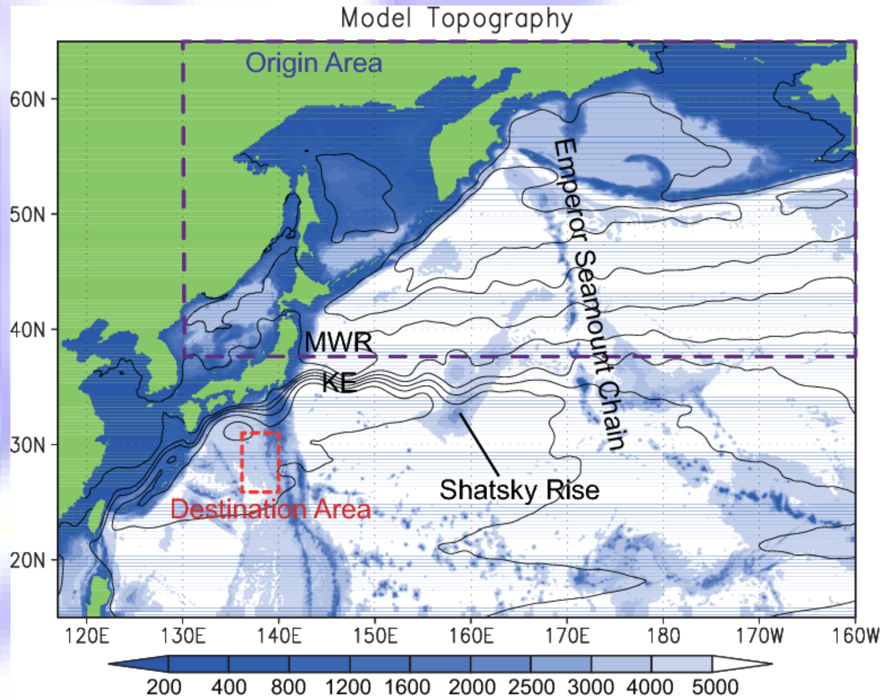


Purpose: Identify the shortcut route of NPIW.

Method: Trace water masses that migrate from the low-salinity (<34.0) near-surface layer in the subarctic region (origin) to the NPIW layer in the sea south of Japan (destination).

Current and diffusion coefficient fields: Free-simulation of an ocean model (MRI.COM, Tsujino et al. 2006) integrated using NCEP-R1 forcing.

Forward model: TL model of the original ocean model with setting density perturbation to 0 at every time step. (Salinity perturbation is employed as a passive tracer.)



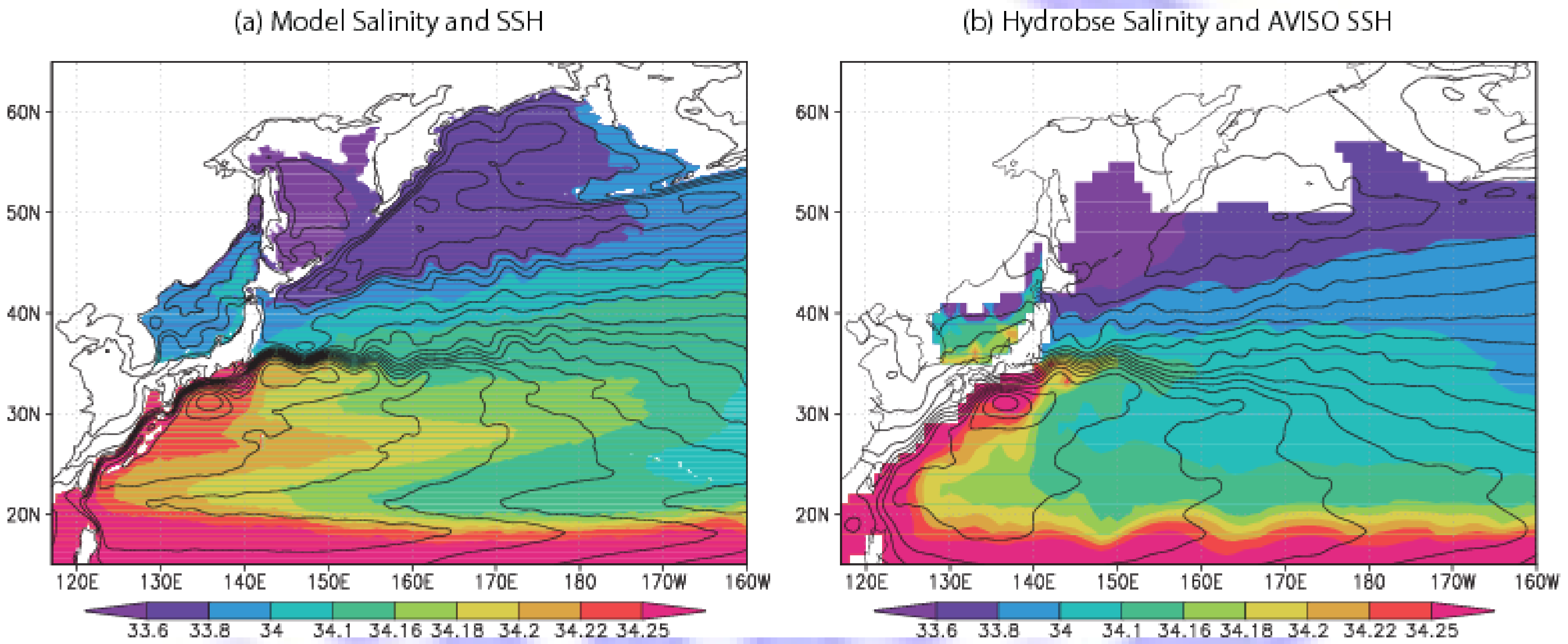
Adjoint model: The exact transpose operator of the TL model

Period of migration (in main analysis):
Jan. 1997 — Dec. 2006 (10-year period)

☆ Background Field (model simulation)

Model

Observation



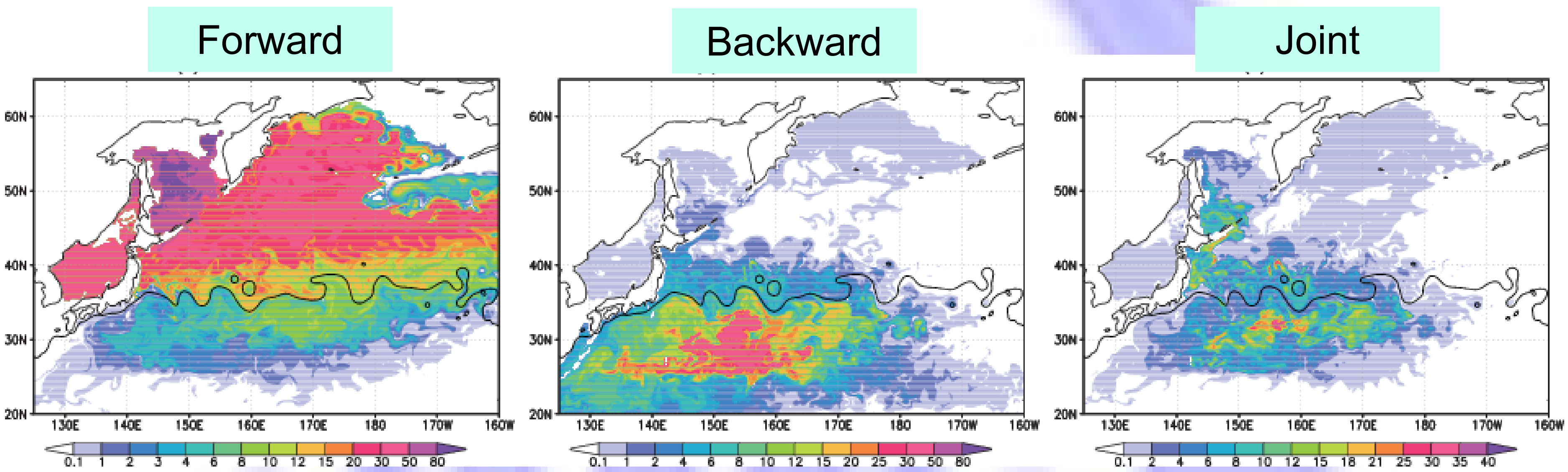
・The first and second ridges of KE and its southward sift between 155-160E is well reproduced although a spurious ridge exists at 155E.

・There are two low salinity tongues at 20N and 30N in the simulation, these are combined into one in the observation.

5 Results

☆ Comparison of Forward, backward, joint likelihood

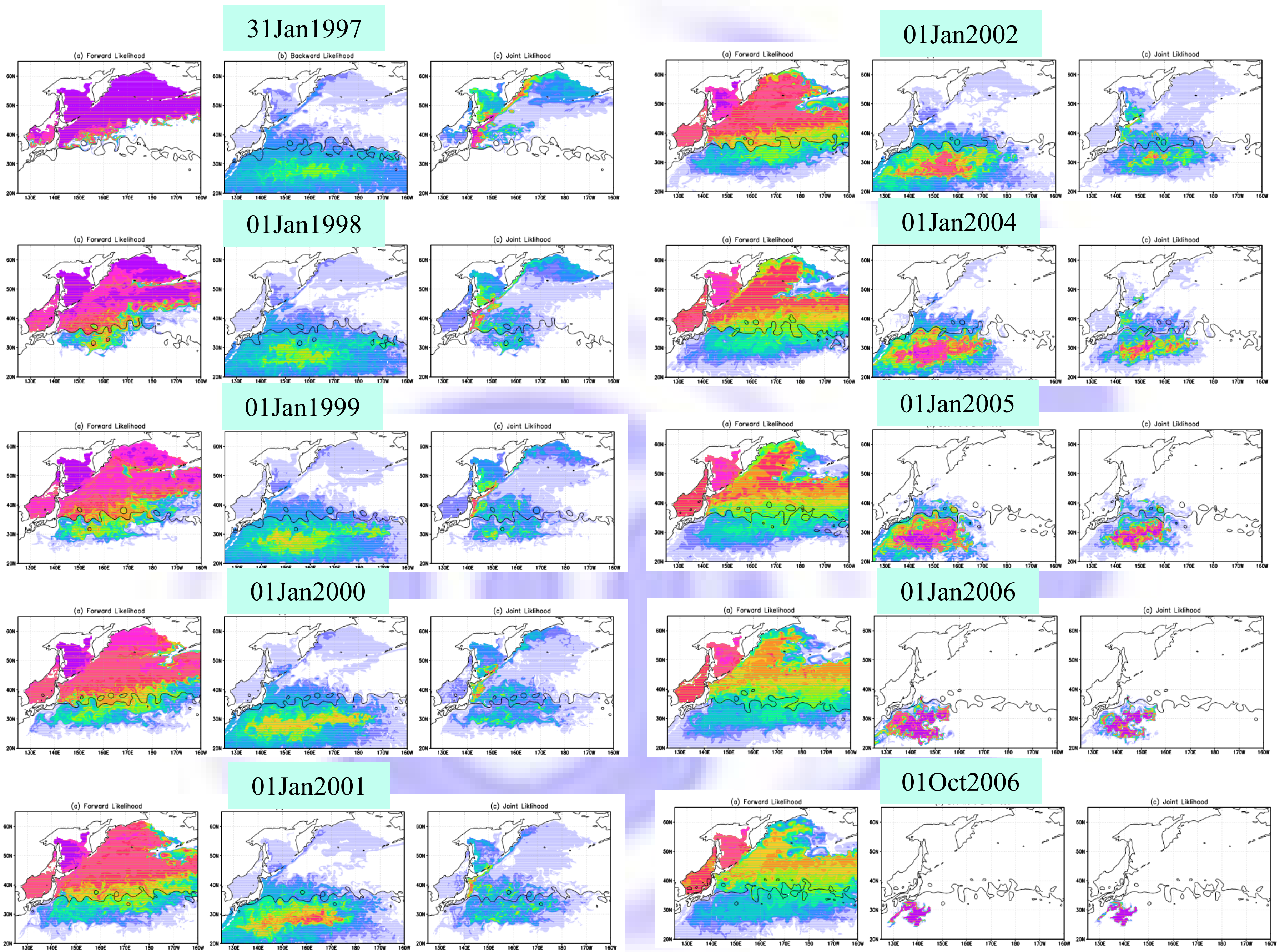
At Jan. 1st, 2002, on the $26.8\sigma_\theta$ surface



・Forward: Particles cross the Kuroshio to the south mainly in 155E-170E.

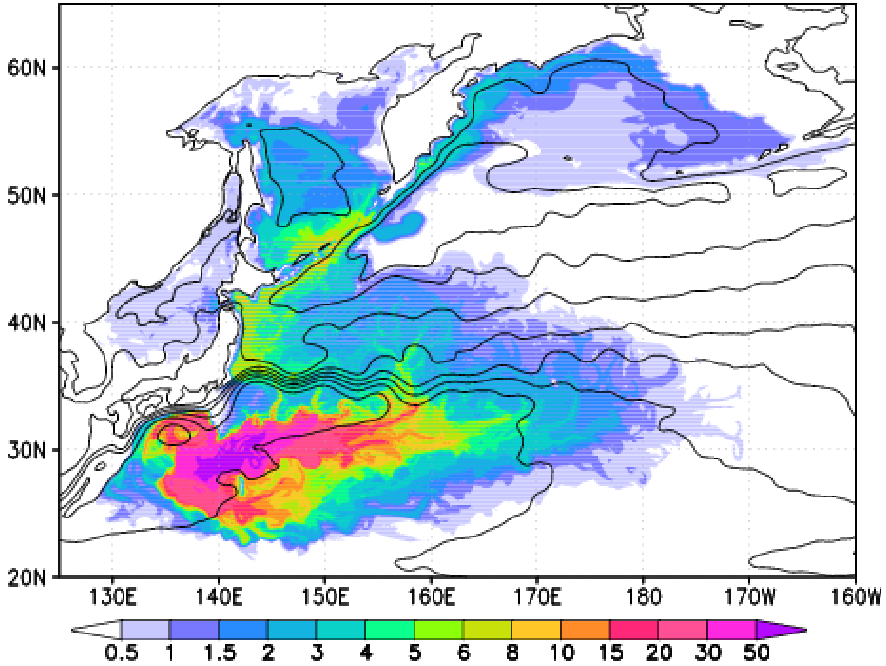
・Adjoint: Although particles are mainly returned to the east in the subtropical gyre, a part of the crosses the Kuroshio and reaches the mixed water region and the Okhotsk Sea.

・Joint: Particles cross the Kuroshio around 160E to the north and reaches the mixed water region and the Okhotsk Sea.

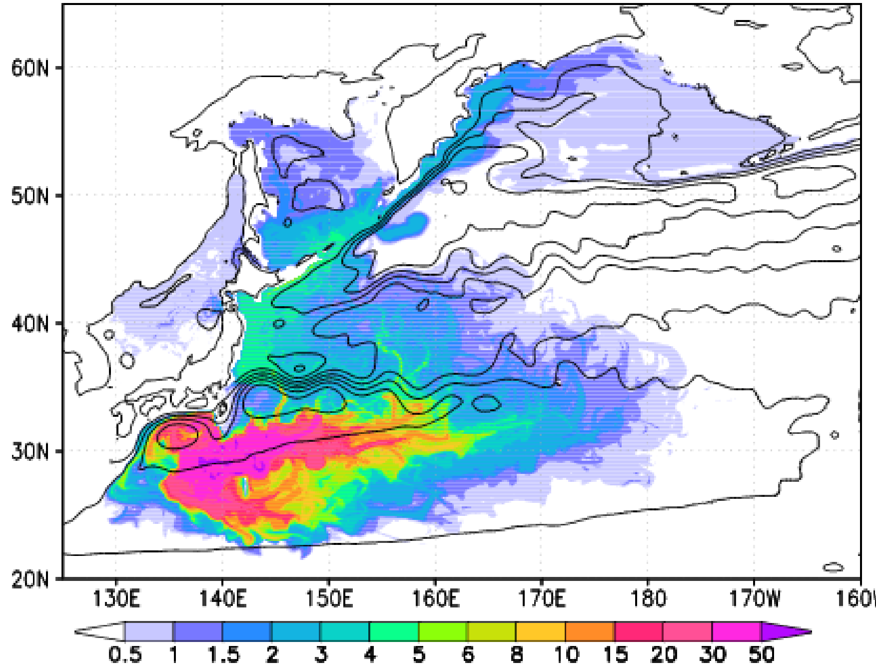

31Jan1997
(a) Forward Likelihood
(b) Backward Likelihood
(c) Joint Liklihood
01Jan1998
01Jan1999
01Jan2000
01Jan2001
01Jan2002
01Jan2004
01Jan2005
01Jan2006
01Oct2006

☆ Maximum joint likelihood in the migration period

Vertically integrated value with SSH



On 26.8σ_θ with acceleration potential



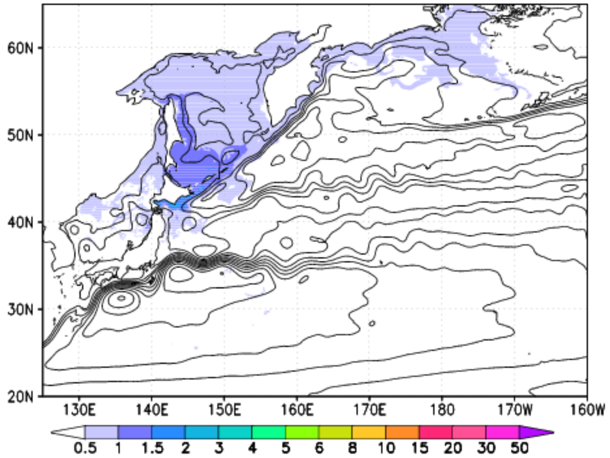
・This figure roughly visualizes the pathways of NPIW.

・The main pathway is advected by Oyashio → 1st ridge of KE → advected by KE → Shatsky Rise → Subtropical gyre → Destination

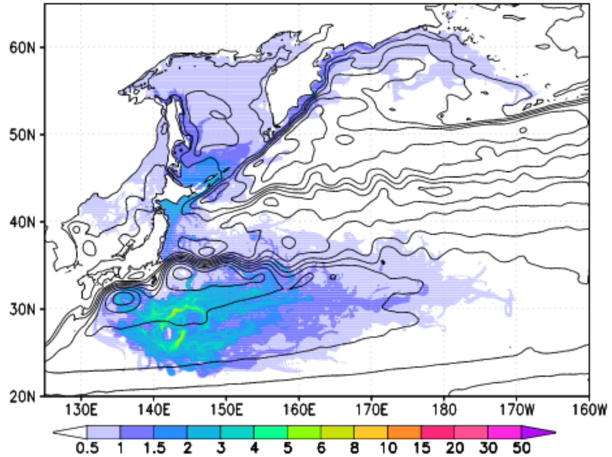
・The pathway north of the Kuroshio is more remarkable in vertically integrated value (because of the effect of the vertical diffusion).

★ Distribution on different density surfaces

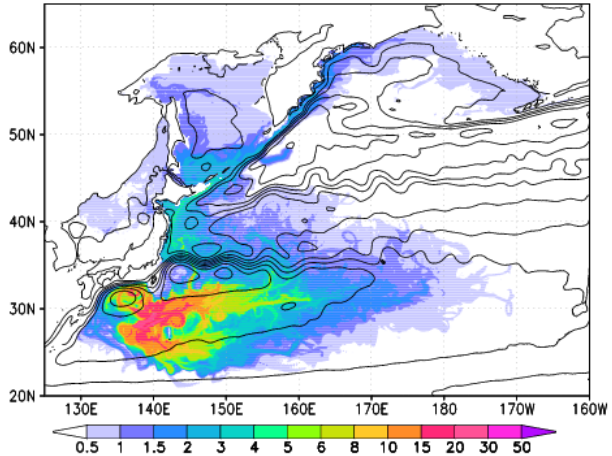
$26.0\sigma_\theta$



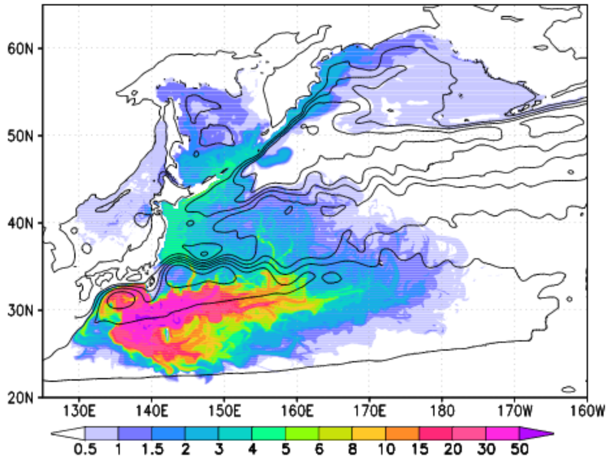
$26.3\sigma_\theta$



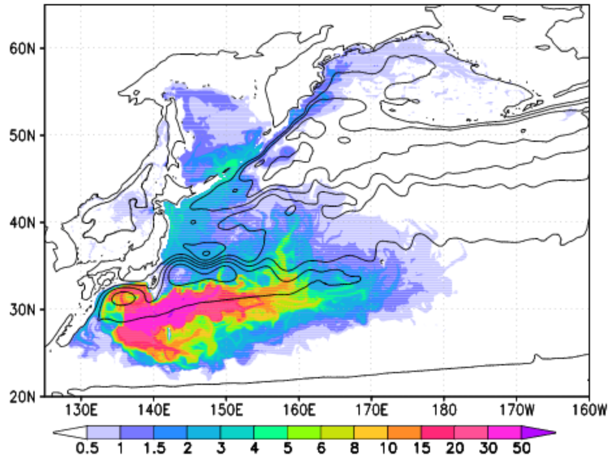
$26.5\sigma_\theta$



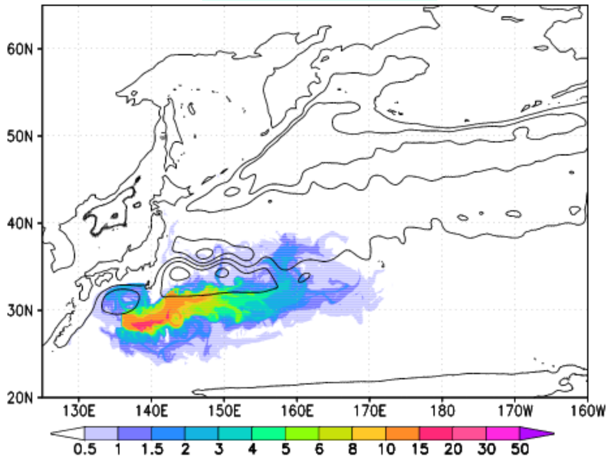
$26.8\sigma_\theta$



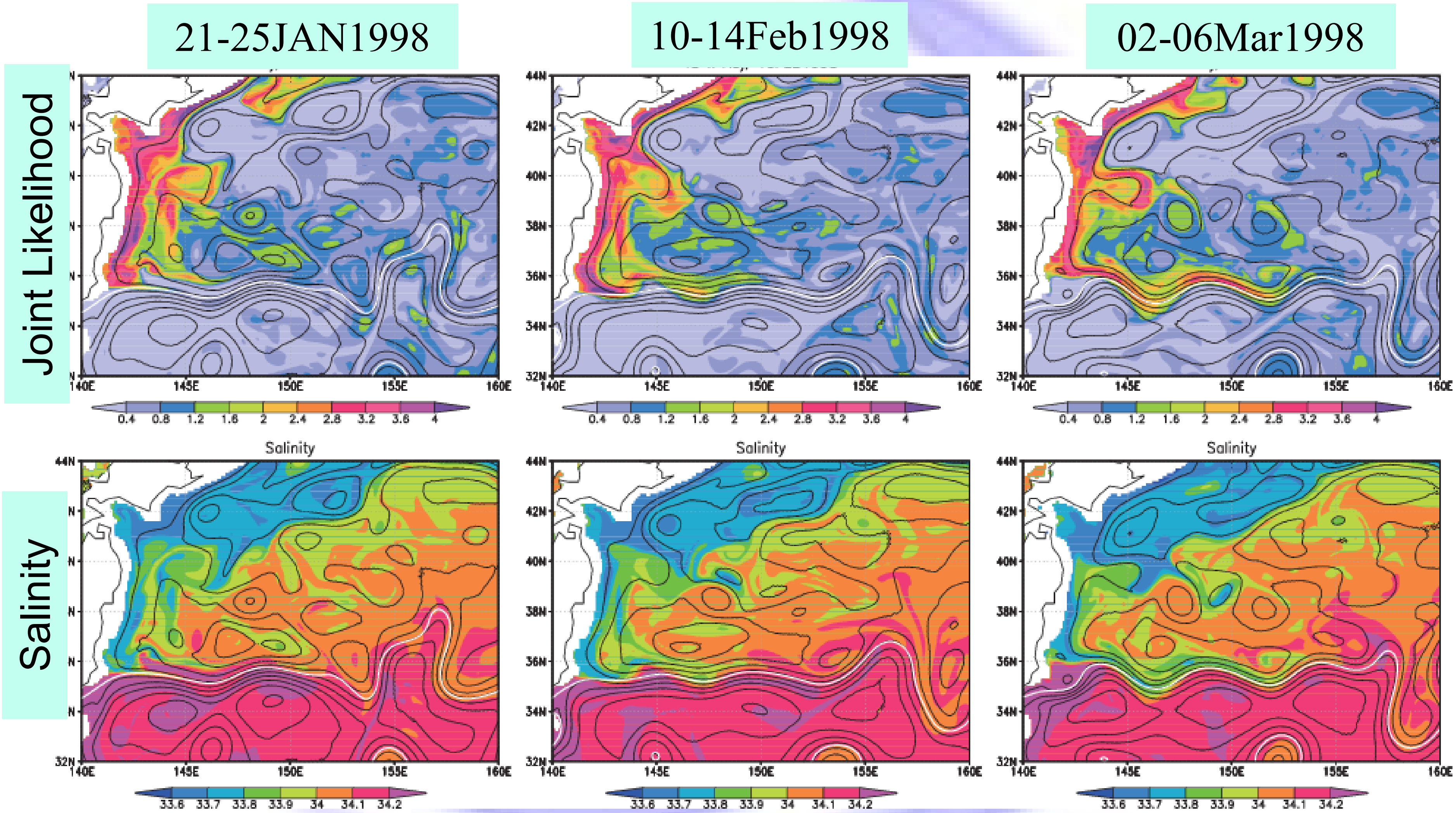
$27.0\sigma_\theta$



$27.2\sigma_\theta$

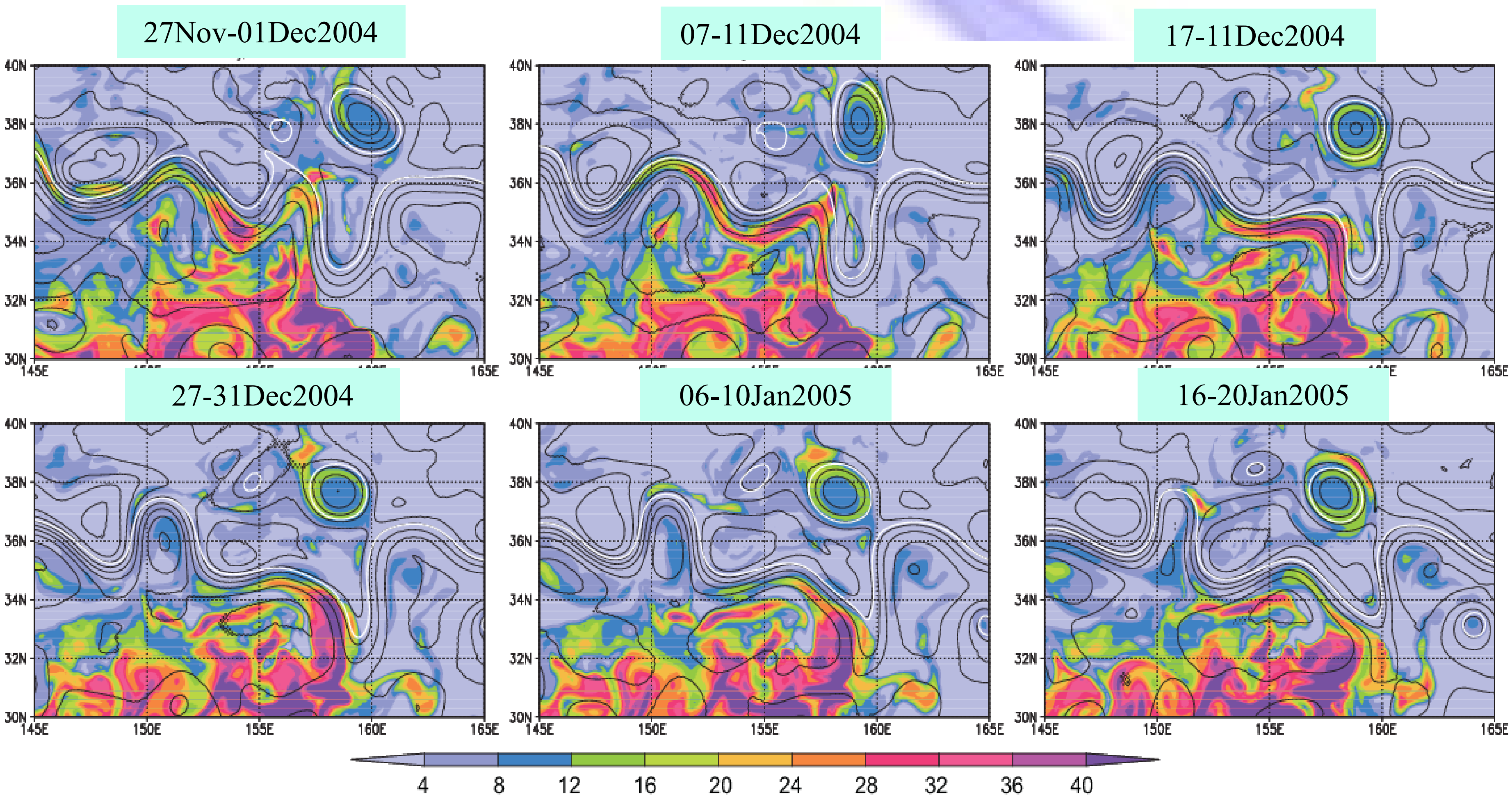


★ Intrusion to the Kuroshio Extension（$26.8\sigma_\theta$）



・Particles that intrude into the Kuroshio and are advected to the east corresponds to the low salinity water.

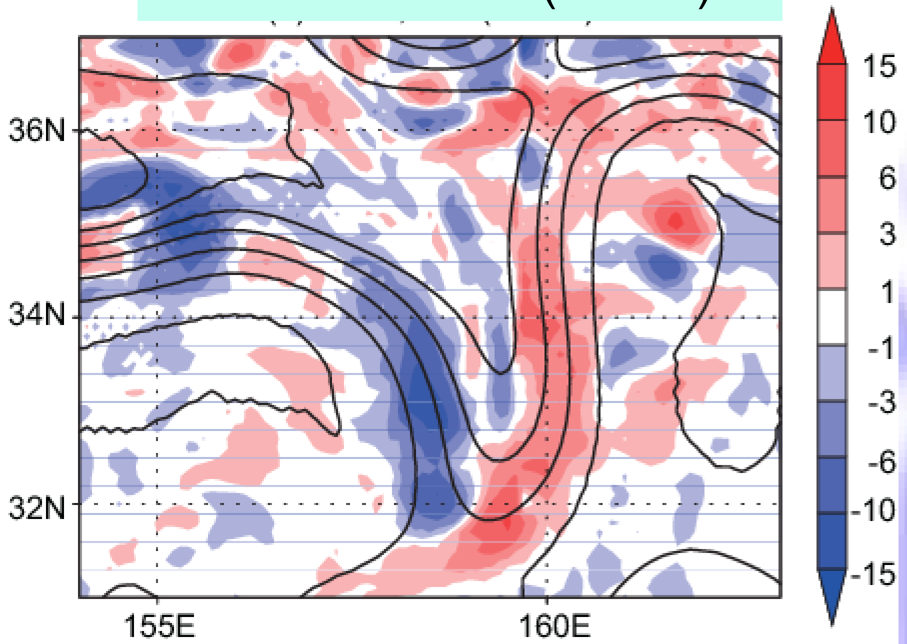
☆ Ejection from the Kuroshio around Shatsky Rise (1)



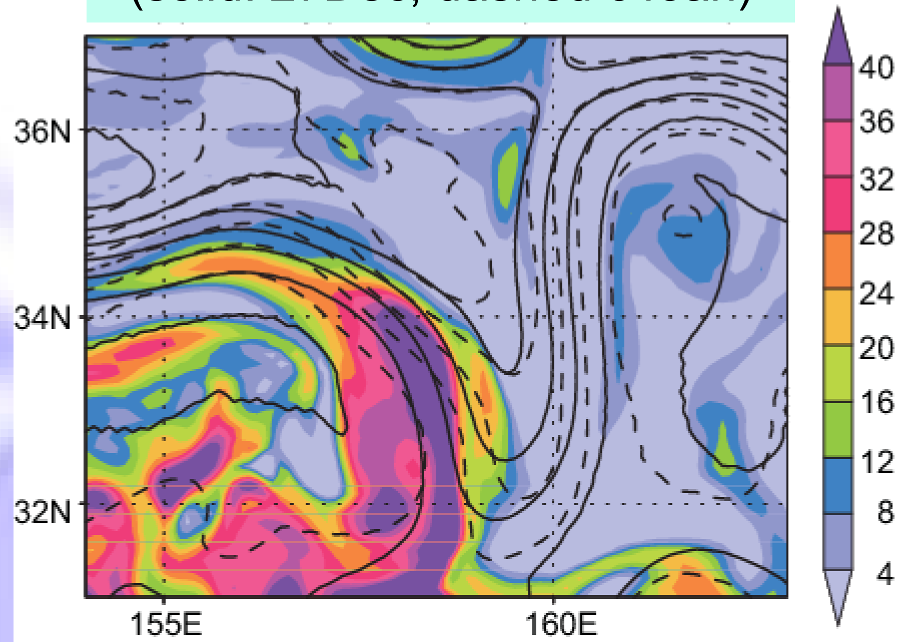
Particles deviate from the Kuroshio around the meander over the Shatsky Rise!?

★ Why do particles are ejected at the meander?

Stretch/shrink of vortex tubes and ACP (27Dec)



Joint likelihood and ACP (solid: 27Dec, dashed 01Jan)

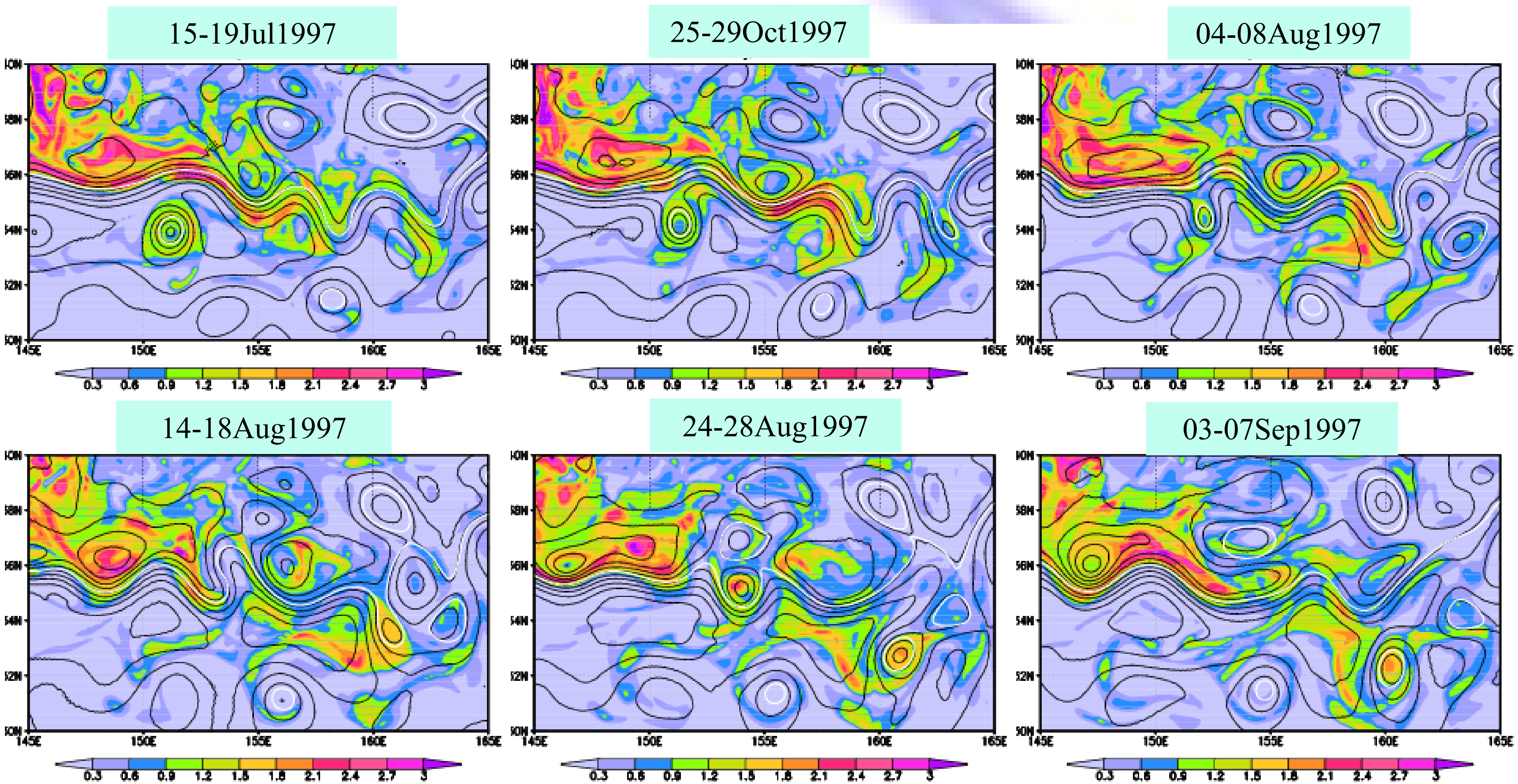


Water mass in the western side of the meander gets anticyclonic vorticity due to the shrinking of the vortex tube.

⟹ Enhance the deviation from the Kuroshio current

⟹ Promote the migration of the meander to the east

☆ Ejection from the Kuroshio around Shatsky Rise (2)



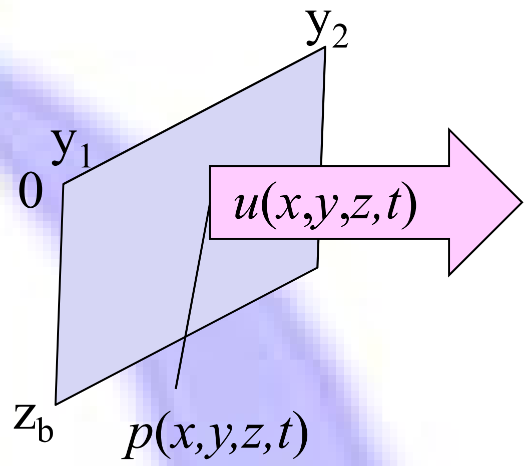
・Some particles enter the subtropical gyre with eddies separated from the meander of the Kuroshio Extension.

★ Probability Transport

$$U = \int_{t_i}^{t_f} \int_{0}^{z_b} \int_{y_1}^{y_2} up(x,y,z,t)\,dydzdt$$

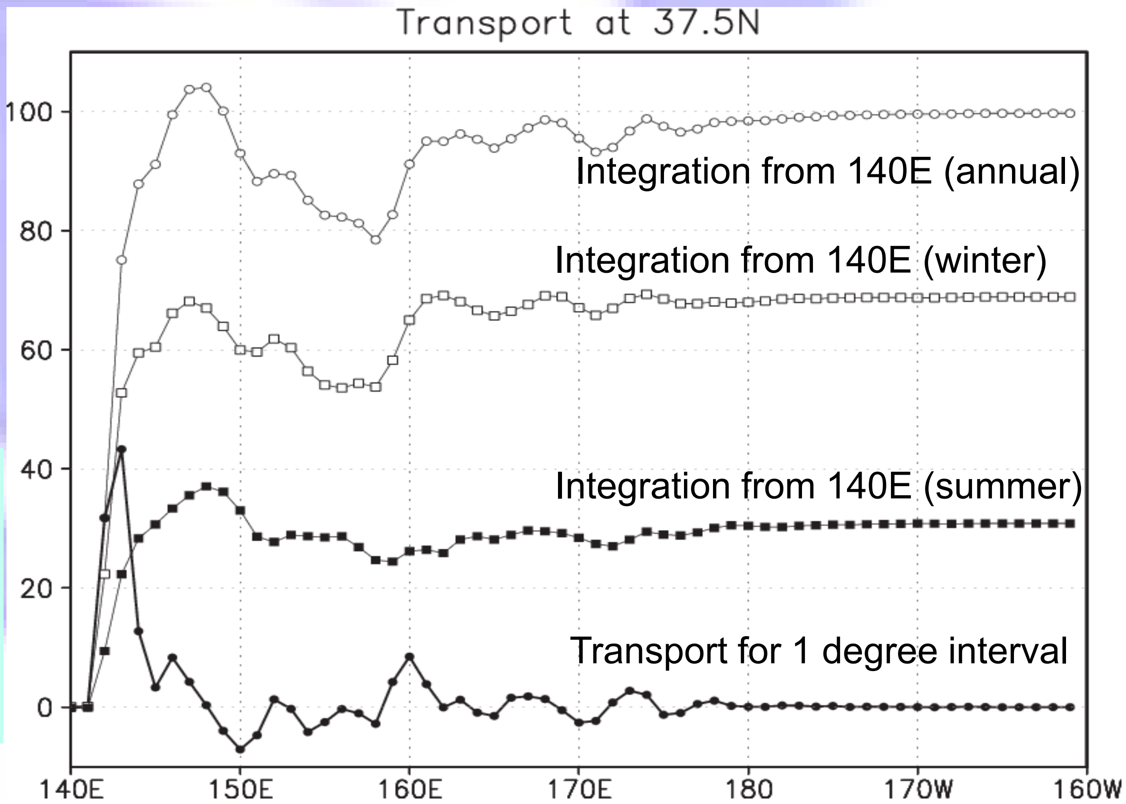
Velocity (m/sec)

Joint PD (% / m^3)



Transport(%):

Fraction of particles that passes the vertical section between y_1-y_2 in the period of t_i - t_f .

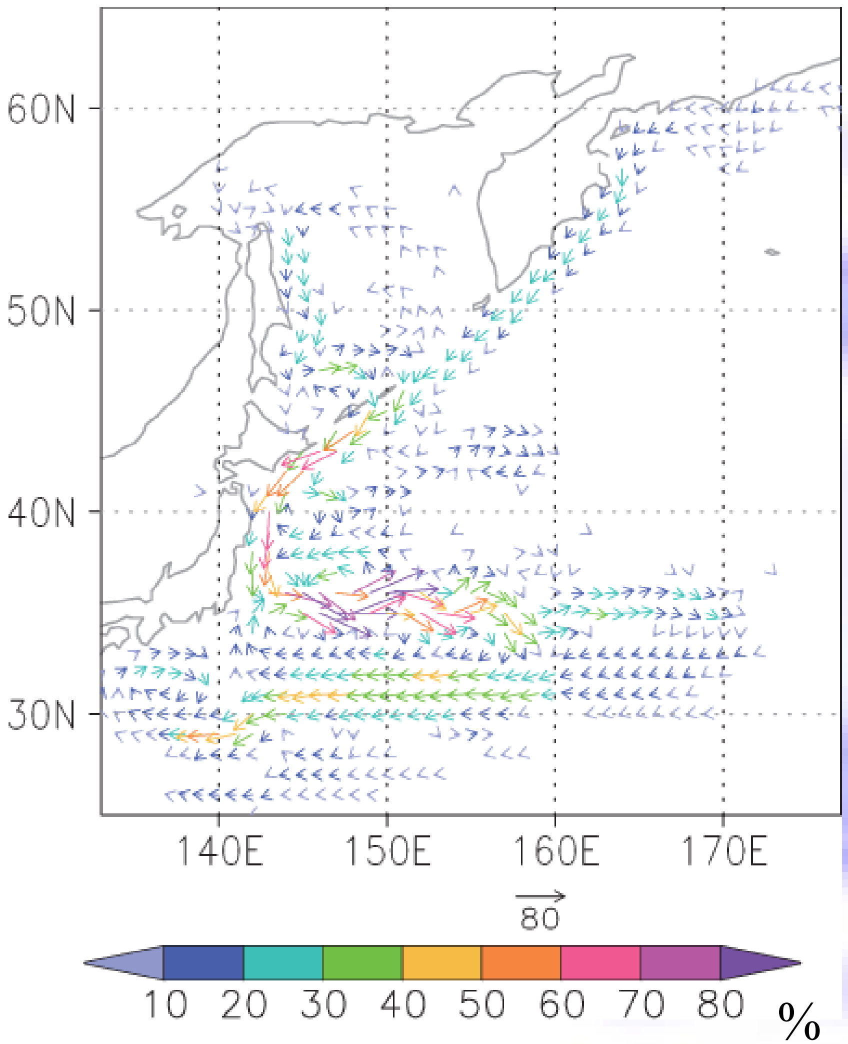


Fraction of particles that crosses 37.5N is almost 100%.

⇨ The effect of diffusion can be ignored.

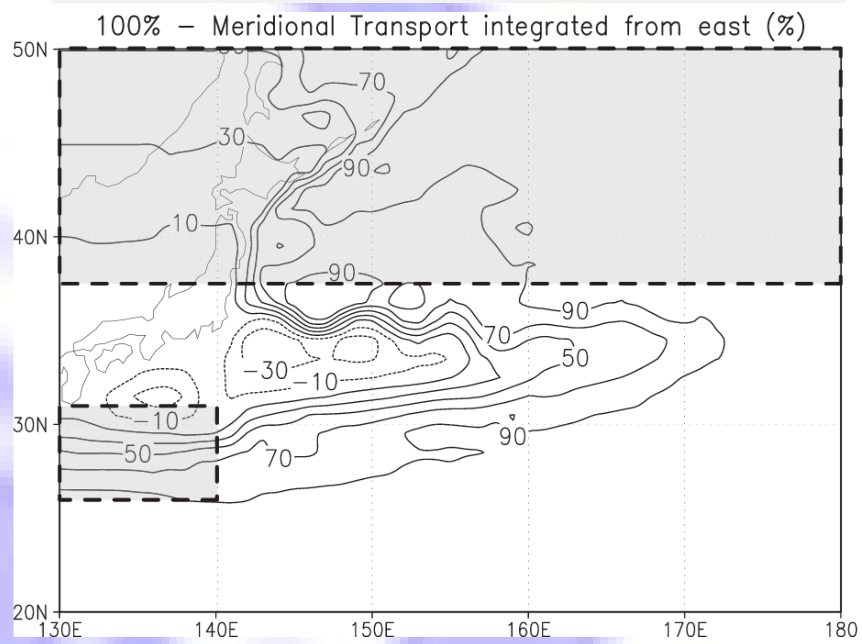
★ Map of the Transport

Transport Vector (integrated over the whole migration period)



100% - Meridional transport integrated from east (%)

Contours can be regarded as stream lines in the area without shading.



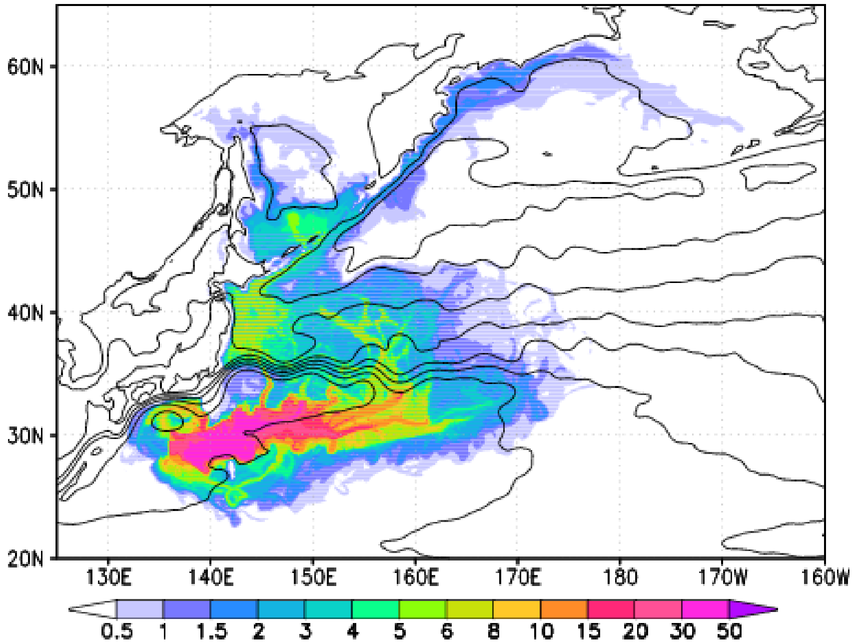
・Okhotsk or Bering Seas → Oyashio → Kuroshio Extension → Subtropical Gyre → Destination

・Particles mainly enter the subtropical gyre over the Shatsky Rise and the west of the Emperor Seamount Chain.

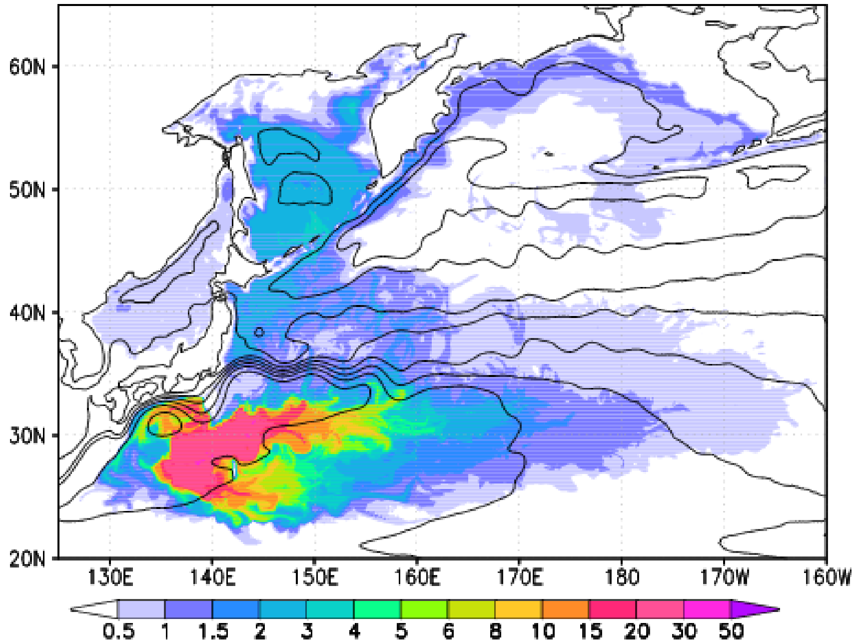
6 Discussions

☆ Dependency on the migration period

2002-2006（5-year period）



1987-2006（20-year period）

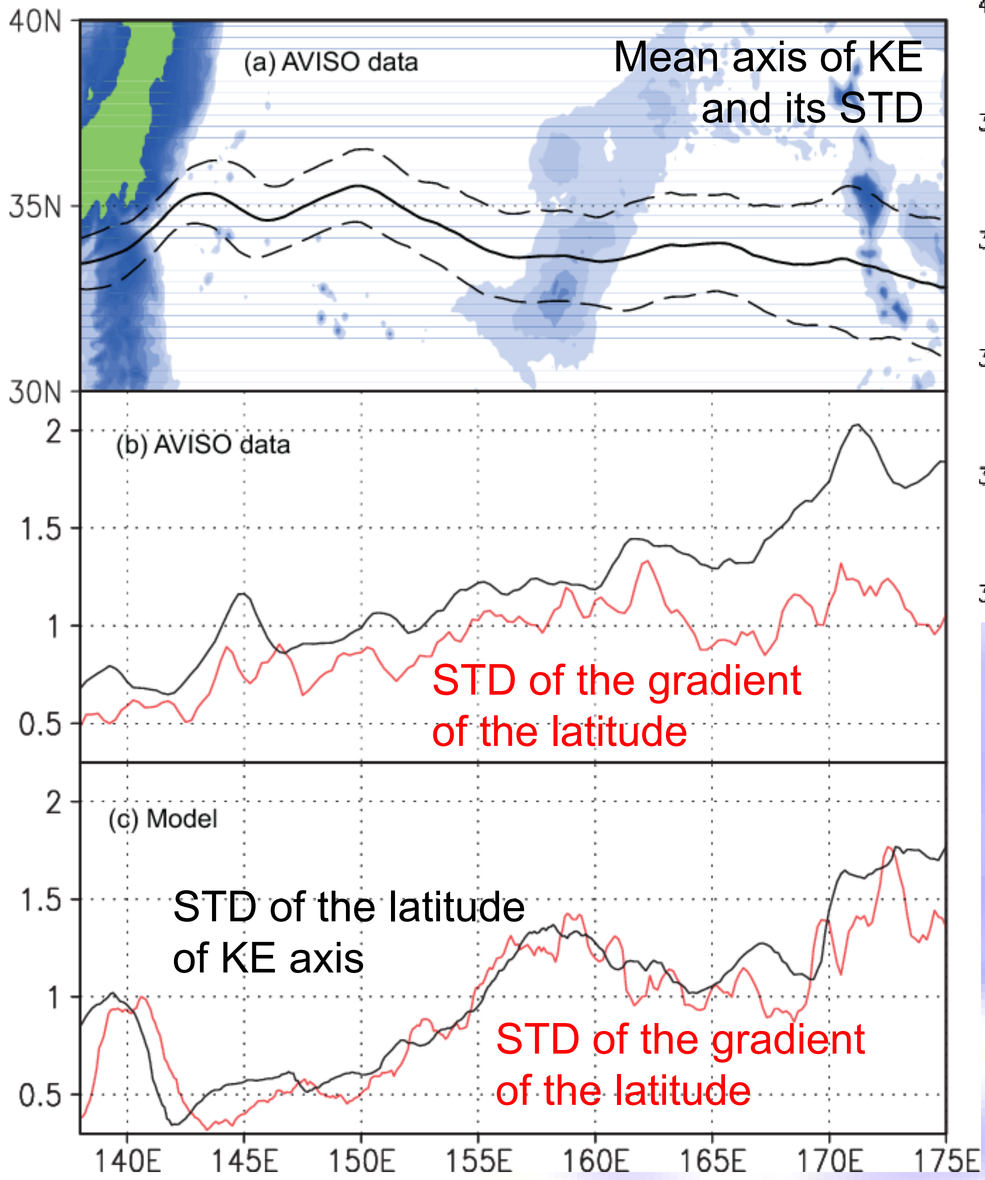


・Pathways is not severely changed with the migration period.

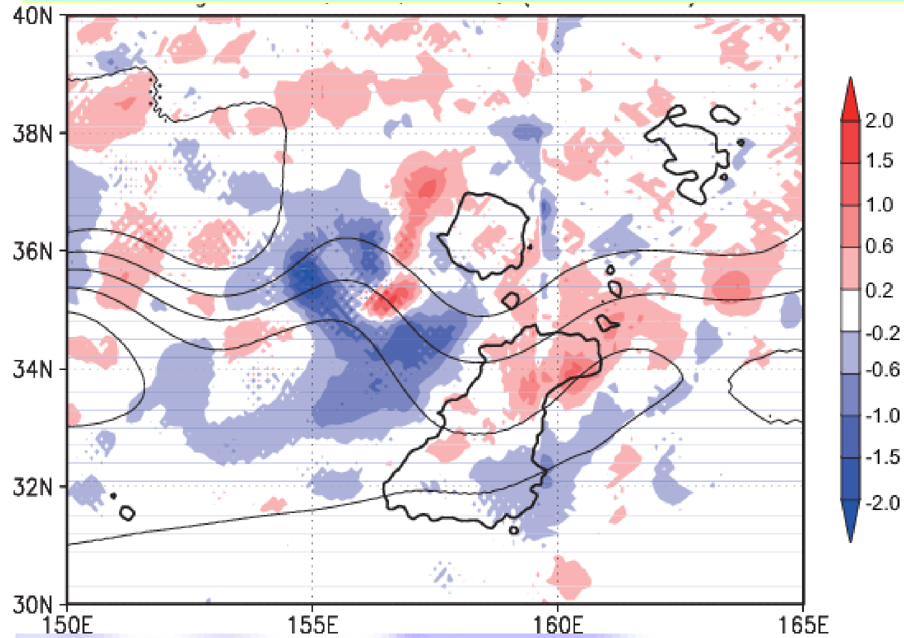
・The longer the period is, the more the pathways become diffusive.

→ The longer periods allow various pathways.

★ Effect of the Shatsky Rise



Mean stretch/shrink of vortex on 26.8σ_θ
(with ACP and 4000m bottom-depth)

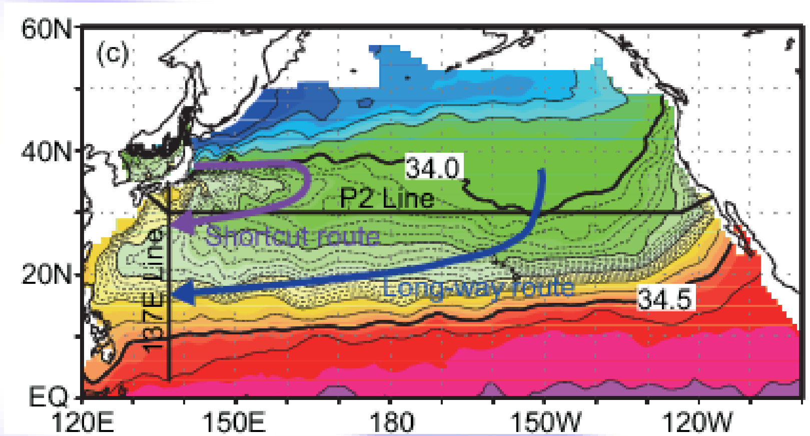


Standard Deviations (STDs) of the latitude of KE axis and its gradient is relatively large east of 155E.

Anticyclonic vorticity tends to be generated at the upstream side of the Shatsky Rise.

☆ On the long-way route

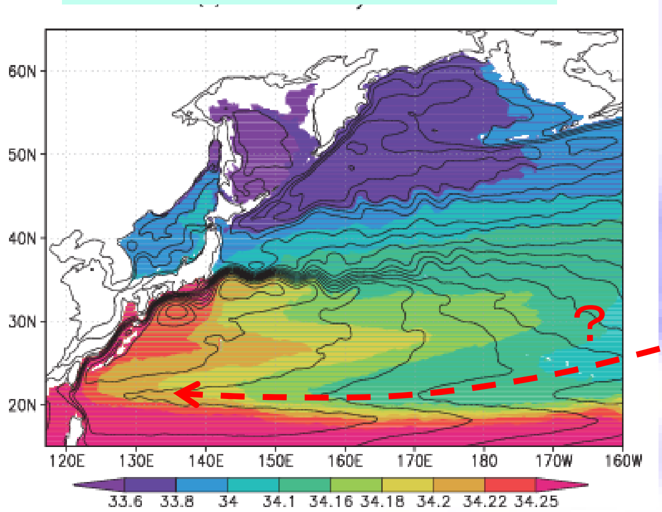
・We hope the shortcut route is identified in this study.



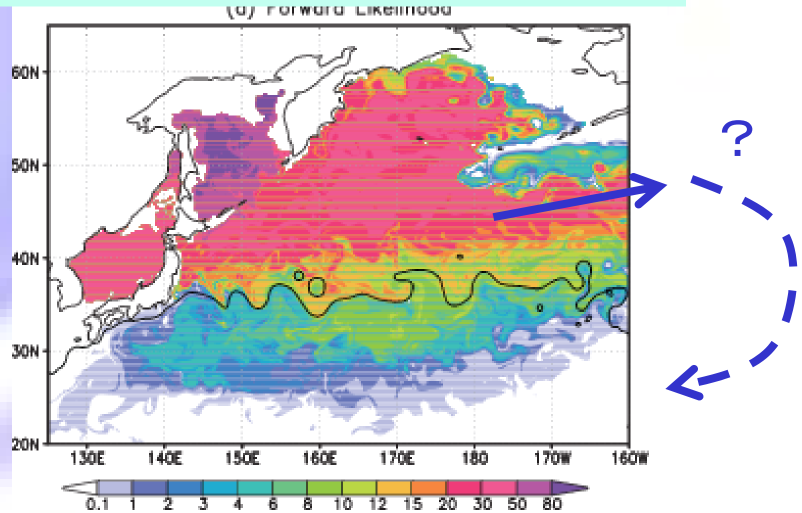
・The long-way route is not identified because it passes through the outside of the ocean model.

・But the simulation result shows some hints on the existence of the long-way route.

Salinity on $26.8\sigma_\theta$



Forward likelihood on $26.8\sigma_\theta$



7 Concluding Remarks

★ Benefits from Adjoint Models

- ✓ An adjoint model is not just the tool for data assimilation, but a propagator of a sensitivity (gradient of any cost function) backward from output to input data.
- ✓ According to the definition of the cost function, the sensitivity (adjoint variable) can represent, information of data misfit (errors), the fraction of particle that reach an certain area at the final time, or the potential impact of the input data on the target physical phenomenon at the final time.

  ⇨ Evaluation of the impact of input data, parameters, observation etc.

- ✓ The connection from the output to the input data represented by the adjoint model is based on the physics in the original forward model (not on the statistical information like with ensemble methods).
- ✓ Conservation of the forward-adjoint inner-product may also have an essential physical mean.
- ✓ Adjoint models (at least partly) hold this property even if the original models have nonlinearity.