1 **GEM: A Dynamic Tracking Model for Mesoscale Eddies in**

2 **the Ocean**

4   Qiu-Yang Li[1], Liang Sun[1, 2], Sheng-Fu Lin[1]

5   [1]School of Earth and Space Sciences, University of Science and Technology of China, 230026, Hefei, China.

6   [2]State Key Laboratory of Satellite Ocean Environment Dynamics, Second Institute of Oceanography, State Oceanic

7   Administration, Hangzhou, 310012, PR China.

8   *Correspondence to*: L. Sun (sunl@ustc.edu.cn)

10   **Abstract**

11   Genealogical Evolution Model (GEM) is an efficient logical model used to track dynamic evolution of mesoscale

12   eddies in the ocean. It can distinguish different dynamic processes (e.g., merging and splitting) within a dynamic

13   evolution pattern, which is difficult to accomplish using other tracking methods. To this end, GEM first uses a two-

14   dimensional (2-D) vector rather than a scalar to measure the similarity between eddies, which effectively solves the

15   ''missing eddy'' problem (temporally lost eddy in tracking). Second, GEM uses both parents and children in tracking,

16   and the dynamic processes are described as birth and death of different generations. Additionally, a look-ahead

17   approach with selection rules effectively simplifies computation and recording. All of the computational steps are

18   linear and do not include iteration. Given the pixel number of the target region L, the maximum number of eddies M,

19   the look-ahead time steps N, and the total number of time steps T, the total computation complexity is O

20    $(LM(N+1)T)$. The tracking of each eddy is very smooth because we require that the snapshots of each eddy on

21   adjacent days overlap one another.

22   Although eddy splitting or merging is ubiquitous in the ocean, there were related different distribution patterns when

23   we applied GEM to track eddies in the Northern Pacific Ocean. Both the merging and splitting rates of the eddies

24   were high, especially at the western boundary, in currents and in "eddy deserts." GEM is useful not only for

25   satellite-based observational data but also for numerical simulation outputs. It is potentially useful for studying

26   dynamic processes in other related fields, e.g., the dynamics of cyclones in meteorology.

28   **1   Introduction**

29   Eddies are ubiquitous in the ocean, and they move from one place to another [Chelton and Schlax, 1996; Chelton et

30   al., 2007]. Eddies in the ocean can cause large-scale transports of heat, salt and other passive tracers [Bennett and

31 White, 1986; Chelton et al., 2011a; Dong et al., 2014; McGillicuddy et al., 2011] by trapping these passive tracers

32 inside the eddies. Such transports has important impacts on the environment and climate of the ocean [Dong et al.,

33 2014]. To address various applications in the studies that use satellite products of sea level anomaly (SLA) data [e.g.,

34 Chelton et al., 2011b] and numerical simulation outputs [e.g., Petersen et al., 2013], oceanic eddies should be

35 automatically recorded using these data and outputs [e.g., Yang et al., 2013; Sun et al., 2014; Pegliasco et al., 2015].

36 In general, the recording of oceanic eddies often includes two independent steps: automated eddy identification and

37 automated eddy tracking. The eddies are identified in a sequence of SLA maps using an identification algorithm. An

38 automated tracking procedure is then applied to determine the trajectory of each eddy [Chelton et al., 2011b].

39 Recently, several automated identification and tracking algorithms have been developed for eddies in the ocean

40 [Chelton et al., 2011b; Ienna et al., 2014; Mason et al., 2014; Yi et al., 2015].

41 For the eddy tracking stage, according to a recent census [Wang et al., 2015; Yi et al., 2015], approximately 10-30%

42 of eddies may be found in proximity to a neighboring eddy in any given global SLA map. Therefore, an eddy

43 tracking process should have the capability to distinguish different dynamic processes (e.g., merging and splitting)

44 during its dynamic evolution. Moreover, an eddy tracking process must be accurate and fast enough to handle a huge

45 amount of data, which will be even larger in size if spatio-temporal resolution of observations and numerical

46 simulations increases.

47 Various implemented automated tracking procedures differ in detail, but they are all similar in concept because they

48 utilize the closest eddy strategy [Chelton et al., 2011b]. For each eddy E_i identified at time step k, the closest eddy to

49 E_i at the next time step $k+1$ is identified as part of the trajectory of eddy E_i. A more advanced procedure uses eddy

50 shape error as an additional condition when searching for an eddy trajectory [Mason et al., 2014].

51 However, there is a "missing eddy'' problem that must be solved in the eddy tracking stage [Chelton et al., 2011b].

52 An eddy at time step k may have no associated eddy at time step $k+1$, which is simply due to a temporary missing

53 eddy in the identification process; this can occur for a variety of reasons related to sampling errors and measurement

54 noises [Chelton et al., 2011b]. Chelton and his colleagues made an attempt to accommodate such problems; they

55 allowed for the reappearance of a temporarily missing eddy by looking ahead two or three time steps. Unfortunately,

56 this "look-ahead" procedure considers too many nearby eddies as potential ones. In practice, the results of this

57 simple ''look-ahead'' procedure were disappointing because the resulting eddy trajectory often jumped from one

58 eddy to another. As a result, it was abandoned, even though this look-ahead feature is highly desirable [Chelton et al.,

59 2011b].

60 Recently, the concept of multiple hypothesis assignment (MHA) was introduced to solve the missing eddy problem

61 by abandoning the simple closest eddy strategy and applying a new "look-ahead" procedure [Faghmous et al., 2013].

62 The MHA method can effectively solve the missing eddy problem in a straight line model when the following

63 trajectory being followed is a branch without any splitting, but it is algorithmically and computationally complex.

64 Given the maximum number of eddies in any time frame M, the number of look-ahead time steps N (with N=0 being

65   the original linear closest eddy procedure without look-ahead) and the total number of time steps T, the MHA has a

66   larger computational complexity, $O(M^{N+1}T)$ [Faghmous et al., 2013].

67   The existing straight-line model can trace the kinematic motion of eddy. The dynamic evolutionary processes (e.g.,

68   merging and splitting) of the eddy are, however, ignored by the model. This implies that each eddy E_i identified at

69   time step k has only one eddy as part of its trajectory at time step k-1 and has only one eddy as part of its trajectory

70   at time step k+1. In the ocean, small eddies may merge to form larger ones. As shown in Figure 1, the anticyclonic

71   eddies AC1 and AC2 observed on July 26, 2006 merged into a single one on July 31, 2006. Then, the cyclonic

72   eddies C1 and C2 on July 26, 2006 merged to form a larger one on August 3, 2006. To describe such processes, the

73   eddy tracking records should be trees with branches instead of simple straight lines.

74   To record the dynamic evolution of eddies, two fundamental algorithms are required. First, the two nearby eddies

75   should be distinguished in the identification stage using a segmentation strategy in which the target region is divided

76   into two corresponding eddies. Otherwise, the merging and splitting processes cannot be determined properly. This

77   problem was recently solved by the use of segmentation strategies, e.g., the close-distance segmentation strategy [Li

78   et al., 2014] and the watershed strategy [Li and Sun, 2015]. Because these segmentation strategies can distinguish

79   closed eddies, they can also potentially reduce the risk of having a missing eddy in the identification process.

80   Second, the merging and splitting processes in the tracking stage should be described in details. We use a multi-

81   branch tree model to do so. The eddy E_i identified at time step k may be more than one eddy at time step k-1, which

82   are merged; and it may become more than one eddy at time step k+1. We refer to this model as the "Genealogical

83   Evolution Model (GEM)" because it is a genealogical tree for recording the whole evolutionary history of an eddy.

84   The multi-way tree model in computer science can be used to store this type of structure.

85   Moreover, the GEM also provides a new way to solve the missing eddy problem. Instead of the existing closest eddy

86   strategy, a temporal track tree with N look-ahead time steps is used to maintain all possible tracks with the help of

87   the multi-way tree model. The method can effectively solve the missing eddy problem, regardless of whether the

88   eddy is splitting or not.

89   In this paper, we introduce the GEM to describe mesoscale eddies in a tracking process with a total number of time

90   steps T. The GEM allows the eddy to have multiple eddies as its parents or as its daughters in a multi-branch model.

91   It also solves the missing eddy problem by using a new look-ahead method similar to the MHA. Compared with the

92   computational complexity $O(M^{N+1}T)$ of MHA, the new method is much faster and has much less computational

93   complexity $O(LM(N+1)T)$, where L the pixels of target region. If the algorithms work well, the output data will

94   record the dynamic evolution of the eddy in detail and will potentially be useful for other research field, e.g., the

95   dynamics of cyclones in meteorology. The GEM is applied to eddies in the North Pacific Ocean (NPO) only, to limit

96   the size of the study area, and we assume the eddies do not cross the equator.

97   The paper is organized as follows. The data and eddy detection method used in this study are introduced in section 2.

98   Then GEM is introduced in section 3, including similarity vector, look-ahead approach and the computation

99   complex. Results including eddy tracks and examples of merging and splitting events are illuminated in section 4.

100 The impacts of data noises and parameters on the results are discussed in section 5. Finally, a summary and
101 conclusions are given in section 6.

102

103 **2   Eddy detection**

104 **2.1   Input data**

105 The input data mainly include the original flow field, which can come from satellite observations or numerical
106 simulations. The flow field used in this study is the 20-year (1993-2012) *daily* SLA data from the merged and
107 gridded satellite product of Maps of Sea Level Anomaly (MSLA) at $0.25^o \times 0.25^o$ resolution in the global ocean by
108 the AVISO (http://www.aviso.oceanobs.com/). The data were corrected for all geophysical errors by the data
109 provider. In this study, we use the "DT14" (delayed-time 2014) altimeter product [Duacs/AVISO, 2014], which is
110 adequate for direct eddy detection [Capet et al., 2014] though it still has about 2-3 cm error globally for short
111 temporal scales [Carrere et al., 2016]. A comprehensive discussion of gridded products for eddy investigations can
112 be found in Chelton et al. (2011b).

113 We used the original SLA data ("DT14") without any filtering or smoothing to identify eddies in this study.
114 However, this does not imply that data smoothing is not needed for the SLA data in related studies. For example, to
115 calculate some eddy parameters (e.g., velocity and vorticity), some kind of smoothing is required to remove data
116 errors, as pointed out by Chelton et al. (2011b). Moreover, the data errors, even if they are very small, might affect
117 eddy detection (see discussion in section 5.1).

118 **2.2   Eddy identification**

119 The eddy identification used in this study is similar to those used before [Chelton et al., 2011b; Mason et al., 2014],
120 to identify eddies in the SLA data. The following mononuclear eddy definition is also similar to what was used by
121 other authors [Chaigneau et al., 2011; Li et al., 2014; Li and Sun, 2015]. Each pixel has eight nearest neighbours. A
122 point within the region is a local extremum if it has an SLA greater or less than all of its nearest neighbours. We also
123 use such definition of extremum in our following analysis, in which the extrema are identified by checking each
124 pixel in the map along with the eight pixels around it. An eddy is defined as a simply-connected set of pixels that
125 satisfies the following criteria:

126   (1) only *one* SLA extremum exists in the pixel set;

127   (2) the SLA value of the eddy is above (below) a given SLA threshold;

128   (3) the amplitude of the eddy (the max difference of SLA values) is larger than a critical value (e.g., 1 cm);

129   (4) the area of the eddy must be large enough (say >16 pixels).

130   Conditions (2)-(4) provide the lower bounds for eddy size and amplitude. These conditions automatically reduce the

131   total number of detected eddies, especially for the small ones. Condition (2) is the same as the first criteria in

132   Chelton et al., (2011b). It is used in consideration of the 2-3 cm of background SLA error [Carrere et al., 2016]; so,

133   small fluctuations in SLA field would not be taken as eddies in this study. Condition (3) was generally used

134   previously [Chaigneau et al., 2011; Chelton et al. 2011]. Condition (4) is more restrictive than the generally used

135   value of eight pixels [e.g., Chelton et al., 2011; Li et al., 2014]; so, this condition is an add-on, which is potentially

136   useful when deriving eddy parameters using a nonlinear optimal fitting method [Wang et al., 2015; Yi et al., 2015].

137   If the eddy area is too small (only a few pixels), its parameters are very sensitive to its area (number of pixels). The

138   above criteria also remove the constraints of eddy pixel number maximum (e.g., <1000) and eddy size (e.g., <400-

139   1200 km) [e.g., Chelton et al. 2011; Mason et al., 2014]. So, they are simpler and more consistent.

140   The SLA extremum so determined is called eddy center. The set of pixels belonging to an individual eddy is referred

141   to the territory of the eddy, and the outmost SLA contour is the boundary of the eddy. We use the territory and

142   boundary to calculate the similarity of eddies in section 3.2.

2.3   Necessity of segmentation

144   Figure 2 illustrates the necessity for eddy segmentation using the merging process of two eddies. Two different

145   mononuclear algorithms are used in the upper and lower rows. In the top panels of Figure 2, eddies are identified by

146   non-segmentation algorithm. Such mononuclear eddies may be very small. The time evolutions from t=1 to t=3

147   show a decay scenario of two closed eddies C1 and C2. Both their amplitudes and areas become smaller and smaller

148   with time. Then, a large eddy C3 suddenly appears in the same region without any premonition. It is hard to see

149   what happened during the time from t=1 to t=3 from the above-identified eddy parameters. In contrast, the bottom

150   panels of Figure 2 show a merging scenario of two closed eddies C1 and C2 using the segmentation algorithm [Li

151   and Sun, 2015]. During the time from t=1 to t=2, both their amplitudes and areas are seldom changed, while their

152   distance is continually shortened. Then, a large eddy C3 naturally emerges in the same region, while C1 and C2

153   disappear. It is recognized from the eddy data that C3 is the merging result of C1 and C2.

2.4   Eddy segmentation

155   Each eddy is identified by the following procedures. First, we find a simply-connected region with a given threshold.

156   Second, we check whether there is at least one extremum in the region. Third, we check whether the region satisfies

157   the eddy conditions (2) and (3). Finally, we check whether the eddy is multinuclear. Since both (2) and (3) allow the

158   eddy to be multinuclear, we explicitly add condition (1) as a constraint. However, we need a segmentation method to

159   implement this.

160   Figure 3 illustrates this eddy segmentation strategy. Figure 3a shows two individual but nearby eddies. The pixels

161   between the two dashed lines are naturally divided by the watershed. As shown in Fig. 3b, the cross section of the

162   eddy clearly shows that two closely located particles P_1 and P_2 on the left and right sides of watershed slide along

163 their ways to different eddy centres. The shape of SLA can provide sufficient information to segment the
164 multinuclear eddy into mononuclear ones.

165 Herein, we use the Mononuclear Eddy Identification (MEI) of the Universal Splitting Technology for Circulations
166 (USTC) with watershed segmentation [Li and Sun, 2015] to produce eddy parameters, including amplitude, radius,
167 territory, and boundary (Fig. 3), which might be potentially used in other studies [Sun et al., 2014].

168 The GEM mainly represents the logical relationship of eddies, which is less dependent on physical parameters. To
169 this end, the GEM takes the previously-identified eddies by MEI (with territory/boundary, see section 2.2) as its
170 input data.

171

172 **3   Dynamic tracking**

173 **3.1   Overview of GEM**

174 The GEM is a logical model used for tracking the dynamic evolution of mesoscale eddies in the ocean (Fig. 4). The
175 model essentially establishes logical relationships of previously-identified eddies. The relationships are determined
176 by two relatively independent steps: first, measuring the "map link," and then connecting to the "track tree."

177 The first part of GEM is "map link," which uses input eddy data (with territory/boundary, see section 2.2) to
178 establish the link between eddies in different snapshots. In this part, we use a 2-D vector rather than a passive scalar
179 to measure the similarity between eddies E_1 and E_2 on two neighboring days (Figs. 5 and 6, see section 3.2 for
180 details). We then use a relatively complex look-ahead procedure to solve the missing eddy problem (section 3.3).
181 This new look-ahead approach has a duration of N days (Fig. 7). Finally, the links between the eddies in different
182 snapshots are saved (see section 3.3 for details).

183 The second part is "track tree," which uses the outputs from "map link" (i.e., eddy links), as its input (Fig. 4). It
184 connects the eddy links from branches to a tree with the genealogical model (Fig. 8) using two sub procedures:
185 "eddy branch" and "eddy tree." In the "eddy branch" part, we use *parent* and *child* to define the eddy relationship
186 and define all possible types of eddy states: birth, death, living, missing, merging and splitting (Fig. 8a).
187 Consequently, we identify different roles in the eddy branches (see section 3.4 for details). Finally, in the "eddy
188 tree" procedure, we connect the branches based on their roles in the genealogical tree (Fig. 8b). The output of GEM
189 includes eddy tracks and the records of eddy relationships (see section 3.5 for details).

190 In short, the GEM uses previously-identified eddies and/or their links to make dynamic tracks via a genealogical tree
191 model. In addition, it includes two parameters, the critical value of area ratio r_c and N, as inputs. See section 5.2 for
192 discussion on the impacts of these parameter choices.

3.2   Similarity vector

193

194   To establish the connection of the previously-identified eddies, the first part of GEM evaluates the similarity of

195   these eddies. As shown in Figure 5a, there were eddies A1, A2 and B1 detected on March 28, 1997. In Figure 5b,

196   there were four eddies, A1, A2, B1, and B2 on March 29, 1997. We overlapped the eddy territories into a single map

197   (Fig. 5c). Then, we used the intersection of eddy territories on different days to calculate the similarity. For eddies

198   A1 and A2, the intersection was very close to both territories on the first day, and on the second day. For eddy B1,

199   the intersection was close to the territory on the second day, but it was only part of that on the first day.

200   Consequently, eddies A1 and A2 had full similarity on these days, while eddies B1 and B2 only had partial

201   similarity on these days.

202   To estimate the above similarity, let us describe it in a logical way. As shown in Figure 6a, there is an eddy (E_1) that

203   is identified by the thick contour of Boundary 1 in the rectangular comparison region on days 0, and there are three

204   eddies (E_2, E_3 and E_4) that are identified in the same region on day 1. This comparison region, which is centered at

205   the eddy center of E_1, moves in time with the target eddy (E_1). To determine the similarities between E_1 on day 0 and

206   E_2 to E_4 on day 1, we overlap the eddy territories into a single map. For example, to determine the similarity

207   between E_1 and E_2, we count the overlap area S_{12} (defined as the intersection of Boundary 1 and Boundary 2)

208   between E_1 (area S_1) and E_2 (area S_2), and then we calculate the following ratios:

209    $$r_1 = S_{12} / S_1 \,,$$ (1a)

210    $$r_2 = S_{12} / S_2.$$ (1b)

211   Clearly, the values of r_1 and r_2 are within [0,1]. The larger r_1 and r_2 are, the larger possibility E_2 has to be the

212   snapshot of E_1 on day 1. Eddy movement speeds are generally less than 0.1 m/s, which implies that an eddy can only

213   move one grid (0.25°) in 3-4 days. Thus, the overlap of the same eddy territory should be large enough. We choose

214   r_c=2/3 for this study, and the choice of r_c is comprehensively addressed in section 5.2.

215   Using the vector (r_1, r_2) and the critical value r_c, we define four different types of similarity between two eddies (Fig.

216   6b). From low to high, they are as follows: Type 0 (T0), where E_1 and E2 are unrelated; Type 1 (T1), where E_1 on

217   day 0 is part of E_2 on day 1 (E_1 enlarging or merging); Type 2 (T2), where E_2 on day 1 is part of E_1 on day 0 (E_1

218   decaying or splitting); and Type 3 (T3), where E_1 and E_2 are the same eddy at different locations on different days

219   (E_1 living and moving). The last type (T3) can also be identified in cases when the center of E_1 propagates less than

220   a pixel toward that of E_2, because the eddy movement speed is physically less than one grid (0.25°) per day. For

221   example, eddy B1 on March 29, 1997 in Figure 5b is simply assigned to T3 even though $r_1 < r_c$. Eventually, we

222   obtain the relationships between E_1 and E_3 or E_4 (Fig. 6a). Because the present method uses a vector to express eddy

223   similarity, we call it the similarity vector. This is an alternative way for dimensionless similarity parameters [e.g.,

224   Ienna et al., 2014; Mason et al., 2014].

225   For example, as shown in Figure 6a, the good similarity between E_1 and E_2 over a critical value r_c (marked as T3 in

226   Fig. 6b) suggests a connection from E_1 to E_2. Eventually, the "eddy branch" procedure (in section 3.4) establishes an

227 inheritance connection in Figure 6a according to the branch definition in Figure 8a. This is similar for eddies E_1 and

228 E_3, but with a different splitting relationship (marked as T2 in Fig. 6b). However, the relationship between eddies E_1

229 and E_4 is designated as "unrelated" because of the overlap in their territories is small or zero. In other words, their

230 overlap rates are below the critical value r_c (marked as T0 in Fig. 6).

231 In previous eddy tracking studies, simple methods were used for weekly SLA data (delayed-time 2010), e.g., the

232 closest distance between eddies [Chelton et al., 2011b; Yi et al., 2015], the closest direction between eddies [Zhang

233 et al., 2014] and the dimensionless similarity scalar [Chaigneau et al., 2008; Mason et al., 2014]. There is always a

234 risk of eddy jumping (from one track to another) in these methods, except for that of Pegliasco et al. (2015), who

235 used intersections of eddy boundaries to find the continuing eddy. Compared to the previous tracking methods, we

236 use a more robust technique to measure neighboring eddies by using the overlap in their territories. In addition, we

237 do not simply assign the continuing eddy using the similarity vector for the two adjacent days; rather, we try to solve

238 the temporary missing eddy problem by looking ahead a few days.

3.3   Look-ahead

240 In contrast to the procedure used in Chelton et al. (2011b), we use a relatively complex look-ahead procedure. Some

241 possibilities for a given eddy are shown in Figure 7a. In the upper row, both E1 and E2 take the same eddy E3 as

242 their subsequent T1 type of eddy, which is a merging event (e.g., eddies C1 and C2 in Fig. 1). Since a T1 eddy has

243 $r_2 < r_c$ (intersection only takes a part of the eddy on day 1), two or more eddies (e.g., Ec1 and Ec2) on day 0 could

244 identify the same eddy (Ec3) as T1 eddy simultaneously on day 1. In the middle row, eddy E1 has two T2 type of

245 eddies (Ec2, Ec3) at the same time; this is a splitting event (e.g., eddies B1 and B2 in Fig. 5). In the lower row, eddy

246 E1 has T2 and T3 types of eddies (respectively Ec2, Ec3) at the same time. Although there may be many

247 possibilities for any given eddy, there is at most one eddy that can be marked as a T1 or T3 eddy on the following

248 day (as $r_1 > r_c$ holds).

249 This new look-ahead approach with N daily is shown in Figure 7b. After finishing the calculation of the following

250 eddies on day 1, we continue to calculate eddies on the following days. At this preparation stage, it is similar but

251 slightly different than the MHA method [Faghmous et al., 2013]. What makes this look-ahead procedure novel and

252 efficient is that we use two simple rules to directly choose only one day's result for the following eddies. Thus, the

253 procedure becomes linear without iteration, and it is much faster than the MHA, as discussed in the subsection on

254 the computation complexity (section 3.6).

255 The two selection rules are: 1) the most similar first, and 2) the closest day first. Rule 1 has priority. We first choose

256 the most similar eddy as the potential one according to their types. According to Figure 6b, T2 type eddy covers

257 only part of the original eddy while T1 eddy covers most part of the original eddy. The similarity from low to high is

258 T2<T1<T3. For example, if there is only one T3 eddy in these days, we choose it as the potential one. However, if

259 there is more than one day with the same type of eddies, we need an additional rule: the closest day first. For

260 example, in the upper row of Figure 7b, there is one T3 eddy on day 1, there is one T3 eddy on day 2, and there are

261 two T2 eddies on day 3. In this case, we choose day 1 as the following day and the T3 eddy as the following E1. In

262 the middle and the lower rows, we choose day 2 and day 3 as the following days and the corresponding T3 eddies as

263 the following E1.

3.4 Eddy branch

265 After having determined the look-ahead day and the following eddy types based on the above process, we can now

266 establish the braches of an eddy from one day to the next. To describe the GEM more precisely, we use *parent* and

267 *child* to identify the different roles that the eddy plays in eddy branches. There are three types of logical

268 relationships used in GEM, as shown in Figure 8a.

269 The upper row shows a successor relationship: an eddy P on day 1 has only one successor (eddy P itself) on day 2.

270 In this case, eddy P is allowed to be missing during day 1 and day 2. Additionally, eddy P will be recorded as death

271 (black circle), if no successor eddy is found after N days.

272 In the middle row, two (or more) eddies merge into one. The first type includes principal and subordinate merging.

273 A principal eddy P_1 and a subordinate eddy P_2 on day 1 merge into a larger eddy P_1 on day 2, whereas P_2 is recorded

274 as death. This occurs when a large eddy meets and merges with a small eddy (e.g., C1 and C2 in Fig. 1). The

275 anticyclonic eddies A1 and A2 in Fig. 11 also experience a similar process (see section 4.2 for details). The second

276 type is coordinated merging. Two (or more) parent eddies P_1 and P_2 merge to produce a new child eddy C, and all of

277 the parent eddies are recorded as death. This is because we could not decide to which of P_1 or P_2 the record of eddy

278 C should be appended. There might be another choice by keeping parent eddies P_1 and P_2 alive and appending the

279 record of eddy C to both eddies. This choice artificially increases total records and leads to other tracking problems;

280 so, we simply abandon it.

281 In the lower row, a parent eddy splits into several child eddies. The first type is principal and subordinate splitting. A

282 parent eddy P splits into an eddy P (itself) and a child eddy C, which is recorded as birth. The second type is

283 coordinated splitting. Two (or more) child eddies are born from the parent eddy P, which is then recorded as death.

3.5 Track tree

285 Finally, the track tree is recorded by connecting the eddy branches (Fig. 8b). In this process, the role that an eddy

286 plays in the track tree should be considered. The first generation is the parent eddy (e.g., P_1), the second generation

287 is the child eddy (e.g., C_1) and the third generation is the grandchild eddy (e.g., G_1). The track tree basically uses the

288 above eddy branches (Fig. 8a). We connect the branches from one time to another to obtain the whole eddy track

289 tree.

290 There are two additional notations. First, an eddy emerging from the same family of eddies (e.g., two siblings C_2 and

291 C_4) will be recorded as a new family member (e.g., eddy C_5). Second, an eddy merging from two different families

292 of eddies (e.g., C_1 and P_2) will be recorded as a new eddy N_1.

293 Although the model could have several generations, we only recorded two generations in this study due to the
294 complexity of the output data structure. However, we can indirectly track other generations using the relationships
295 between them.

3.6 Computation complexity

297 To calculate similarity vectors, we need to overlap two small regions around eddy E1. The total number of pixels in
298 the rectangular comparison region is L. The computational complexity of the similarity vector is $O(L)$ for each day.
299 If we use N look-ahead time steps to find the best choice, the computation complexity of the branches will be O
300 $(L(N+1))$ for one eddy. Because all of the steps are linear without iteration, given the maximum number of eddies in
301 any time frame M, the number of look-ahead time steps N and the total number of time steps T, the total computation
302 complexity is $O(LM(N+1)T)$. This might be the fastest method possible. When the number of look-ahead time steps
303 N is more than one, the computation complexity is much faster than $O(M^{N+1}T)$ of MHA.

304 For example, both L and M are approximately 1000, and $N=2$ is used in the present study. The MHA method will
305 require on the order of 10^2-10^3 times more computational time than the present method; and the larger the value of N,
306 the more efficient the present method is. The look-ahead time N may be potentially as large as one week ($N=6$), as
307 noted in the following discussion. Thus, the present method is especially effective when a long look-ahead time is
308 required for poorly identified eddies.

309

4 Results

4.1 Eddy tracks

312 We first apply the MEI to detect the ocean eddies in the North Pacific Ocean (NPO) during 1993-2012. The eddy
313 centers (SLA extrema of eddy snapshots) on each day are counted on each $1^o \times 1^o$ grid. In general, anticyclonic eddies
314 are significantly more than cyclonic eddies. As shown in Figure 9a, the cyclonic eddies are mainly located in the
315 western part of the NPO. For example, there are lots of cyclonic eddies east of Japan near the Kuroshio, which can
316 also be seen from both Figure 1 and the results in section 5.1. In contrast, anticyclonic eddies are mainly located in
317 the eastern part of the NPO (Fig. 9b). For example, the eddies are mainly anticyclones in the red box, which can also
318 be seen from the results in section 4.2. In general, the eddies are ubiquitous in Figure 9c (about 50-70 eddies per
319 year on each $1^o \times 1^o$ grid), except that there are several regions where both types of eddies are relatively scarce. One
320 of them is known as "eddy desert" [Chelton et al., 2007]. The other region is the North Equatorial Countercurrent
321 (NECC) [Hu et al., 2015]. Finally, we present in Figure 9d the ratio of difference in number of cyclonic and
322 anticyclonic eddies to the total number of eddies.

323 We apply the GEM to these eddies detected by MEI with r_c=2/3 and N=2. In the NPO, there are a total of 60276
324 eddies with lifetimes longer than 30 days. Among them, 37553 of the eddies are anticyclonic and 22723 are cyclonic.
325 The tracks of long-lived eddies are plotted in Figure 10. In general, they are similar to those shown in previous

326   studies [Chelton et al., 2011b]. There are 7290 anticyclonic and 3627 cyclonic eddies with lifetimes longer than 100

327   days (Fig. 10a), and the ratio of anticyclonic to cyclonic eddies is approximately 2. The ratio is larger for eddy

328   lifetimes greater than 400 days, which was also noted in previous studies [Chelton et al., 2011b; Xu et al., 2011].

329   Each track is very smooth because we require that the snapshots of eddies on different days overlap one another.

330   The long-lifetime eddy trajectories imply that the quality of the tracking results is reasonable. We will take the long-

331   life C1 in Figure 10b as an example.

332   Eddy C1 was first detected as an eddy initiated on September 14, 1995, with an extremum at 163.5°W, 10.5°N. It

333   then travelled to the northwest and disappeared at 151.25°W, 20.5°N on March 11, 1997. Its trajectory is the longest

334   that we have detected in the NPO (Fig. 10b). The trajectory is smooth, except for a sudden jump from 167.5°E to

335   166.75°E (Fig. 10c) on July 31, 1996. Therefore, a puzzle emerged regarding whether we should connect the

336   trajectories from before July 30, 1996 with that after July 31, 1996, into a single trajectory.

337   To clarify this, we plot the two SLA fields in Figure 10d. The SLA field on July 30, 1996 is plotted as contours. The

338   eddy center is marked by a black cross at 167.5°E, 16.5°N. In contrast, the SLA field on July 31, 1996 is plotted in

339   shading. The eddy center is marked by a red cross at 166.75°E, 17.25°N. The distance between the eddy extrema was

340   larger than 100 km within a day. Although that distance is far beyond the criterion in eddy tracking used in recent

341   studies [Mason et al., 2014; Yi et al., 2015], we can see from the SLA fields that they both indicated the same eddy,

342   and we should connect the trajectories into a single trajectory.

343   A similarity vector can effectively reduce the possibility of temporarily missing an eddy in tracking. In turn, it

344   reduces the usage of the look-ahead procedure. It is clear that the similarity expressed as a vector is better than that

345   as scale using simple distance.

346   **4.2   Eddy merging and splitting**

347   The trajectories also provide evidence of dynamic evolution. The time evolution of a couple of anticyclonic eddies is

348   depicted in Figure 11a, which implies a merging process occurring in the red boxes in Fig. 9. As shown in Figure

349   11a, eddy A1 had a westward movement with a speed of 2.6 cm/s, and eddy A2 lingered near 133°W. Then, both

350   eddies merged into one large eddy on April 23, 1997. That evolutionary process is clearly shown by the SLA fields

351   (Figs. 11c-j). In Figure 11c, there were two anticyclonic eddies, A1 and A2, located at 132°W, 28.5°N. Eddy A1

352   moved from east to west with a nearly constant speed of 2.6 cm/s, whereas eddy A2 had negligible zonal motion.

353   They then rotated clockwise about each other with an average angular velocity of 6×10^{-7} s^{-1}, as denoted by the blue

354   arrows. Finally, they merged into the new large eddy A2.

355   The SLA field shows that an eddy splitting process also occurred. The time evolutions of anticyclonic eddies B1, B2

356   and B3 are depicted in Figure 11b. At first, eddy B1 had a fast westward speed of 10.4 cm/s. It then split into two

357   eddies (B1 and B2) on March 29, 1997 (Fig. 6). Eddy B1 traveled at its original speed whereas eddy B2 lingered at

358   its origin. Then, eddy B3 emerged at a location between B1 and B2 on April 9, 1997, which slowed down the speed

359   of B1 to approximately 3.5 cm/s. After that, eddies B2 and B3 merged into a new eddy B3 on April 19, 1997. In fact,

360 similar to eddies A1 and A2, eddies B1 and B2 eventually merged into a new eddy on May 20, 1997 (not shown).

361 The SLA maps in Figures 11c-j show more details that were not recorded by the eddy tracking data. Note that eddy

362 B2 had a very short lifetime of 20 days but a complex dynamic process. If only long-term eddies (lifetime > 30 days)

363 were saved, the corresponding evolution process might not be recorded properly.

364 It is expected that cyclonic eddies will have a counter-clockwise rotation in the Northern Hemisphere, which is

365 known as the Fujiwhara effect for cyclones [Fujiwhara, 1921]. When two cyclones are close enough, they will begin

366 to orbit cyclonically (counter-clockwise in the Northern Hemisphere). Because the above-mentioned eddies are

367 anticyclonic, they have opposing directions of rotation, which appear as two point vortices moving in circular paths

368 about the center of vorticity in classical fluid dynamics [Batchelor, 1967].

369 **4.3   Census of merging and splitting events**

370 To illuminate how often the merging and splitting processes occurred, we counted the total number of merging and

371 splitting events on each $1°\times1°$ grid each year. The merging and splitting events were ubiquitous in the oceans, but in

372 general were very few times each year per $1°\times1°$ grid element. The merging frequencies for cyclonic eddies and

373 anticyclonic eddies are shown in Figure 12, which are similar to their splitting frequencies (not shown). The

374 distribution pattern of merging frequency for cyclonic eddies in Figure 12a, is very similar to that of cyclonic eddy

375 centers in Figure 9a. In contrast, the merging frequency for anticyclonic eddies was larger along the west coast (Fig.

376 12b), whereas the anticyclonic eddy centers were located mainly in the east (Fig. 9b). Although merging and

377 splitting events may be ubiquitous in the ocean (Fig 12c,d), there are several types of special regions where merging

378 and splitting events occur more frequently.

379 The first type of special region is the western boundary. It is known that the western boundary is a sink of eddy

380 energy caused by the interaction with the bottom and lateral topography [Zhai et al., 2010]. It is also known as a

381 "graveyard" for westward-propagating ocean eddies [Zhai et al., 2010; Chelton et al., 2011b]. The second type of

382 special region is located in strong currents, including the Kuroshio Current, the North Equatorial Current (NEC) and

383 the NECC [Hu et al., 2015]. Among those currents, the eddies in the NECC had the highest frequency of merging

384 and splitting events, which was seldom noted in previous studies. The third type of special region is located in the

385 northeast Pacific, which is also known as an "eddy desert" [Chelton et al., 2007]. The fourth type of special region is

386 located in enclosed marginal seas, especially the Bering Sea.

387 By comparing Figure 12 with Figure 10, we can see that the regions with high frequencies of merging and splitting

388 events have few eddy tracks, especially in the NECC and in the "eddy desert" in the northeast Pacific. The existence

389 of "eddy desert" may be due to the fact that the eddy was too small to be detected or the fact that the eddy lifetime

390 was too short [Chelton et al., 2011b]. However, Figures 9 and 12 suggest that fewer eddies accompanied by frequent

391 dynamic (merging and splitting) events caused the "eddy desert."

392

5   Discussion

5.1   Data noises

Although "DT14" is much better than previous products, there are still some notable errors, especially for short temporal scales of less than two months [Carrere et al., 2016]. It was reported that there are along-track SLA errors of about 2-3 cm globally and of more than 3 cm at high latitudes and in shallow waters.

To reduce the noises in SLA data, one may use the Gaussian structure filter [Chelton et al., 2011b; Mason et al., 2014], Hanning filters [Penven et al., 2005], or Lanczos filter [Chaigneau et al., 2008]. As some parameters are used in these filters, the filtered results depend much on these parameters [see Fig. A1 in Chelton et al., 2011b]. Alternatively, we can simply use a five-point quadratic smoothing to remove the noises in SLA data. The filtered data are then piecewise C^2-smooth, which satisfies the potential requirements for calculating vorticity from SLA data.

Figure 13 shows the non-smoothed and smoothed SLA data from January 1, 1993 to January 4, 1993. The smoothed SLA maps are very close to the non-smoothed SLA maps. And the values at the SLA extrema are close to their original values. These imply that the noises in DT14 data are very small.

However, the noises cannot be neglected, even when they are small. They might induce false SLA extrema (see the definition of extremum in section 2.2), which eventually affect eddy detection, e.g., the false extremum on January 2, 1993 in box A and the false extremum on January 3, 1993 in box B (Figure 13). These false extrema existed only for a very short period (one or two days). But they can induce false merging and splitting events, which may cause eddies to unexpectedly terminate [Chelton et al., 2011b]. This is one of the reasons why we need "look-ahead" in eddy tracking.

5.2   Impact of parameters

To discuss the impact of these parameters, we apply the GEM to the eddies detected in the NPO. There are two parameters in the GEM: the critical value r_c and the look-ahead time N. The numbers of eddies with lifetimes > 30 days are counted for different r_c and N, as shown in Figure 14a. Note that the results are very similar, except for N=0 (i.e., without any look-ahead). It is from the above discussion that we see look-ahead is necessary when there are extrema due to small noises in the data. The numbers of eddies seldom change with r_c for any N>1, when r_c is within 0.5 to 0.8. Meanwhile, the numbers of merging and splitting events are also counted for different r_c and N, as shown in Figure 14b. In general, the splitting events are a little more than the merging events. Note also that the results are very similar, except for N=0. The numbers of merging and splitting events seem to converge for r_c >0.5 as N increases. For each N>0, the numbers of merging and splitting events reach a maximum at r_c=0.6. A relatively loose similarity condition (r_c<0.5) will lead to a risk of eddy jumping from one track to another, which consequently reduces both total eddy number and dynamic events. On the other hand, a relatively strict similarity condition (r_c>0.9) will lead to a risk of missing eddies, which may also reduce both total eddy numbers and dynamic events.

426   In general, the tracking results should be insensitive to the choice of these parameters. From Figure 14, we can

427   observe that $0.5 < r_c < 0.8$ is a potential choice. The optimal value for r_c might be 0.6-0.7. We also find that the look-

428   ahead time N should be larger than 0; otherwise, the risks of eddy jumping and eddy missing are too great. The look-

429   ahead approach effectively reduces such risks. For example, N=1 and N=2 have 95.5% and 98% of the total eddies

430   for N=4, respectively. To reduce the missing eddies to 1%, the look-ahead time might be greater than six days. This

431   is also the physical requirement of the representative period of the merged SLA data [Chelton et al., 2011b].

432   Although N=4 and N=6 might be better, N=2 produced a very similar result (~2% bias) and with a significantly

433   lower computational cost. Our present parameters are reasonable for these regimes.

434   Besides, as noted in section 4.2, there are short-term eddies (lifetime < 30 days), which might to though complex

435   evolution process. If only long-term eddies (lifetime > 30 days) were saved, the corresponding evolution process

436   might not be recorded properly. This should be noted in further applications on eddy dynamics with satellite

437   altimetry data.

5.3   Impact of eddy boundary

439   Eddy boundary is influenced by the identification method and corresponding criteria used. In general, the automated

440   eddy detection algorithms are categorized into three types: 1) physical parameter-based algorithms, e.g., Okubo–

441   Weiss (O–W) parameter [Isern-Fontanet et al., 2003; Chaigneau et al., 2008]; 2) flow geometry-based algorithms

442   [Chaigneau et al., 2011; Chelton et al., 2011b; Wang et al., 2015]; and 3) hybrid methods, which involve physical

443   parameters and flow geometry characteristics [Nencioli et al., 2010; Xiu et al., 2010; Dong et al., 2011; Yi et al.,

444   2015]. For example, Yi et al. (2015) used the O–W parameter to identify eddy kernels and SLA contour geometries

445   to identify eddy boundaries. Thus, it is difficult to directly compare the influences of eddy territory using different

446   tracking algorithms.

447   We can, however, estimate the influence of eddy boundary using an indirect way. Because the eddy center is

448   relatively robust, different identification methods mainly give different eddy boundaries. Consequently, the eddy

449   area S is most sensitive to such an eddy territory. However, the area ratio reduces the sensitivity to the eddy area S

450   because both the overlap area S_{12} and the eddy area S change synchronously. Moreover, our tracking results

451   fortunately are not very sensitive to r_c (or the eddy area S), as noted in the above discussion. For example, the

452   present results are based on a very strict identification method. If we modify the threshold of eddy amplitude from 1

453   cm to 3 cm, the number of identified eddies will decline. However, the identification results for the long-lived eddies

454   appear to be similar (Table 1).

5.4   Future research

456   The GEM is a flexible model that can easily work with other relevant programs, e.g., data filtering and smoothing

457   algorithms [Chelton et al., 2011b; Ienna et al., 2014; Wang et al., 2014], other hybrid eddy detection algorithms [e.g.,

458   Yi et al., 2015] and O-W parameter detection [e.g., Petersen et al., 2013], because the GEM requires a flow field and

459   previously-identified eddies to accomplish dynamic tracking. In addition, the similarity measurement can be
460   replaced by similar methods [e.g., Pegliasco et al., 2015] when considering more complex conditions.

461   The GEM is a complex model. The output data include eddy tracks, relationships and previously-identified eddy
462   parameters (e.g., amplitude and radius). These eddy parameters, which were directly obtained from the identification
463   process, are useful for censuses [Chelton et al., 2011b]. However, they may not be sufficient for some applications.
464   For example, eddy territory was required in our recent studies on typhoons and oceanic eddy interactions [Sun et al.,
465   2010, 2012, 2014]. A better way to obtain these parameters might be to use a nonlinear fitting of the flow field
466   [Wang et al., 2015; Yi et al., 2015] with appropriate models [e.g., Sun, 2011; Zhang et al., 2013].

467   Another future research direction may involve comparing different tracking datasets. Because there are several
468   tracking datasets produced by various methods, it is useful to inter-compare them. This may improve both the
469   tracking methods and the available datasets for further studies.

470   The GEM can be easily applied to larger datasets, even to 3-D numerical simulation outputs [Petersen et al., 2013;
471   Woodring et al., 2016], because its computational time increases linearly as a function of the size of the dataset. The
472   computation of the 20-year daily global SLA data only required a few hours on a personal computer. Such a model,
473   accompanied by other (e.g., velocity-based, O-W-based) identification methods, can be used to analyze numerical
474   simulation outputs.

475   The GEM opens a window to investigate eddy dynamics [Wang et al., 2015] and other applications [Sun et al.,
476   2014]. As illuminated in Figure 11, the dynamic evolution of eddies is accompanied by abundant phenomena that
477   might be identified using the GEM. The present study is only the beginning of such applications.

478

6   Conclusions

480   We have introduced the GEM for tracking dynamic evolution of mesoscale eddies in the ocean. Several novel
481   approaches (e.g., vector similarity and look-ahead approach) were applied to deal with unsolved problems in
482   tracking. All of the computational steps in GEM are linear and do not require iteration. Given the grid number of the
483   target region L, the maximum number of eddies M, the number of look-ahead time steps N, and the total time steps T,
484   the total computational complexity is of $O(LM(N+1)T)$. We applied the GEM to the eddies in the NPO. Each eddy
485   track was very smooth because we required that the snapshots of eddies on neighboring days overlap one another.
486   Both merging and splitting rates of eddies were high, especially at the western boundary, in strong currents and in
487   "eddy deserts." The GEM is useful not only for satellite-based observation data but also for numerical simulation
488   outputs. It potentially has many applications for studies of dynamic processes in related fields, e.g., the dynamics of
489   cyclones in meteorology. The "MEI" and "GEM" computer codes will be provided on request after publication of
490   this paper.

491

492

493 **Acknowledgements**

494 We thank the anonymous referees for their comments and suggestions. We thank the AVISO for providing the SLA

495 data (http://www.aviso.oceanobs.com/). This work was supported by the National Basic Research Program of China

496 (Nos. 2012CB417402 and 2013CB430303), the National Foundation of Natural Science (No. 41376017) and the

497 Open Fund of the State Key Laboratory of Satellite Ocean Environment Dynamics (No. SOED1501).

498

499 **References**

500 Batchelor, G. K. (2000). An introduction to fluid dynamics. Cambridge university press, 615pp.

501 Bennett, A. F., & White, W. B. (1986). Eddy heat flux in the subtropical North Pacific. J. Phys. Oceanogr., 16(4),
502 728-740.

503 Capet, A., E. Mason, V. Rossi, C. Troupin, Y. Fauge `re, I. Pujol, and A. Pascual, (2014), Implications of refined
504 altimetry on estimates of mesoscale activity and eddy-driven offshore transport in the Eastern Boundary Upwelling
505 Systems, Geophys. Res. Lett., 41, 7602–7610, doi:10.1002/2014GL061770.

506 Carrere, L., Faugère, Y., and Ablain, M. (2016). Major improvement of altimetry sea level estimations using
507 pressure-derived corrections based on ERA-Interim atmospheric reanalysis, Ocean Sci., 12, 825-842,
508 doi:10.5194/os-12-825-2016.

509 Chaigneau, A., Gizolme, A., and Grados, C. (2008). Mesoscale eddies off Peru in altimeter records: identification
510 algorithms and eddy spatio-temporal patterns. Progr. Oceanogr, 79, 106–119.

511 Chaigneau, A., Le Texier, M., Eldin, G., Grados, C., and Pizarro, O. (2011). Vertical structure of mesoscale eddies
512 in the eastern South Pacific Ocean: A composite analysis from altimetry and Argo profiling floats, J. Geophys. Res.:
513 Oceans, 116. C11025, doi:10.1029/2011JC007134.

514 Chelton, D.B., Schlax, M.G. (1996). Global observations of oceanic Rossby waves. Science 272, 234–238.

515 Chelton, D. B., Schlax, M. G., Samelson, R. M., & de Szoeke, R. A. (2007). Global observations of large oceanic
516 eddies. Geophys. Res. Lett., 34(15), L15606,. doi:10.1029/2007GL030812.

517 Chelton, D. B., Gaube, P., Schlax, M. G., Early, J. J., & Samelson, R. M. (2011a). The influence of nonlinear
518 mesoscale eddies on near-surface oceanic chlorophyll. Science, 334(6054), 328-332.

519 Chelton, D. B., Schlax, M. G., & Samelson, R. M. (2011b). Global observations of nonlinear mesoscale eddies.
520 Progr. Oceanogr, 91(2), 167-216.

521 Dong, C., Nencioli, F., Liu, Y., & McWilliams, J. C. (2011). An automated approach to detect oceanic eddies from
522 satellite remotely sensed sea surface temperature data. Geoscience and Remote Sensing Letters, IEEE, 8(6), 1055-
523 1059.

524 Dong, C., McWilliams, J. C., Liu, Y., & Chen, D. (2014). Global heat and salt transports by eddy movement. Nature
525 communications, 5:3294, DOI: 10.1038/ncomms4294.

526 Duacs/AVISO (2014), A new version of SSALTO/Duacs products available in April 2014. Version 1.1, CNES.
527 [Available at http://www.aviso.altimetry.fr/fileadmin/documents/data/duacs/ Duacs2014.pdf.].

528 Fang, F., and Morrow, R. (2003). Evolution, movement and decay of warm-core Leeuwin Current eddies. Deep-Sea
529 Res., Part II 50, 2245–2261.

530   Fujiwhara, S. (1921). "The natural tendency towards symmetry of motion and its application as a principle in
531   meteorology". Q. J. R. Met. S. 47 (200): 287–293. doi:10.1002/qj.49704720010.

532   Faghmous, J. H., Uluyol, M., Styles, L., Le, M., Mithal, V., Boriah, S., & Kumar, V. (2013). Multiple Hypothesis
533   Object Tracking For Unsupervised Self-Learning: An Ocean Eddy Tracking Application. In AAAI.

534   Hu, D., Wu, L., Cai, W., Gupta, A. S., Ganachaud, A., Qiu, B., ... & Wang, G. (2015). Pacific western boundary
535   currents and their roles in climate. Nature, 522(7556), 299-308.

536   Ienna, F., Jo, Y. H., & Yan, X. H. (2014). A new method for tracking Meddies by satellite altimetry. Journal of
537   Atmospheric and Oceanic Technology, 31(6), 1434-1445.

538   Isern-Fontanet, J., García-Ladona, E., & Font, J. (2003). Identification of marine eddies from altimetric maps.
539   Journal of Atmospheric and Oceanic Technology, 20(5), 772-778.

540   Li, Q. Y., Sun, L., Liu, S. S., Xian, T., & Yan, Y. F. (2014). A new mononuclear eddy identification method with
541   simple splitting strategies. Remote Sensing Letters, 5(1), 65-72. doi:10.1080/2150704X.2013.872814.

542   Li, Q. Y., & Sun, L. (2015). Technical Note: Watershed strategy for oceanic mesoscale eddy splitting. Ocean
543   Science, 11(2), 269-273. doi: 10.5194/os-11-269-2015.

544   Mason, E., Pascual, A., & McWilliams, J. C. (2014). A new sea surface height–based code for oceanic mesoscale
545   eddy tracking. Journal of Atmospheric and Oceanic Technology, 31(5), 1181-1188.

546   McGillicuddy, D. J. (2011). Eddies masquerade as planetary waves. science, 334(6054), 318-319. doi:
547   10.1126/science.1208892.

548   Nencioli, F., Dong, C., Dickey, T., Washburn, L., & McWilliams, J. C. (2010). A vector geometry-based eddy
549   detection algorithm and its application to a high-resolution numerical model product and high-frequency radar
550   surface velocities in the Southern California Bight. Journal of Atmospheric and Oceanic Technology, 27(3), 564-579.

551   Penven, P., Echevin, V., Pasapera, J., Colas, F., and Tam, J.: Average circulation, seasonal cycle, and mesoscale
552   dynamics of the Peru Current System: A modeling approach, J. Geophys. Res., 110, C10021,
553   doi:10.1029/2005jc002945, 2005.

554   Petersen, M. R., Williams, S. J., Maltrud, M. E., Hecht, M. W., & Hamann, B. (2013). A three-dimensional eddy
555   census of a high-resolution global ocean simulation. Journal of Geophysical Research: Oceans, 118(4), 1759-1774.

556   Pegliasco, C., A. Chaigneau, and R. Morrow, (2015). Main eddy vertical structures observed in the four major
557   Eastern Boundary Upwelling Systems, J. Geophys. Res. Oceans, 120, 6008–6033, doi:10.1002/2015JC010950.

558   Sun L, Yang Y.-J., Xian T., Lu Z. and Fu Y.-F., (2010). Strong enhancement of chlorophyll a concentration by a
559   weak typhoon, Mar. Ecol. Prog. Ser., 404, 39-50, doi: 10.3354/meps08477.

560   Sun, L. (2011). A typhoon-like vortex solution of incompressible 3D inviscid flow. Theoretical and Applied
561   Mechanics Letters, 1(4), 042003.

562 Sun, L., Yang, Y.-J., Xian, T., Wang, Y., and Fu, Y.-F., (2012). Ocean responses to Typhoon Namtheun explored
563 with Argo floats and multiplatform satellites. Atmos. Ocean. 50(sup1), 15-26.

564 Sun, L, Y.-X. Li, Y.-J. Yang, Q. Wu, X.-T. Chen, Q.-Y. Li, Y.-B. Li, & T. Xian (2014). Effects of super typhoons
565 on cyclonic ocean eddies in the western North Pacific: A satellite data-based evaluation between 2000 and 2008, J.
566 Geophys. Res. Oceans, 119(9): 5585–5598, doi:10.1002/2013JC009575.

567 Wang, R., Yang, Z., Liu, L., Deng, J., & Chen, F. (2014). Decoupling noise and features via weighted L1-analysis
568 compressed sensing. ACM Transactions on Graphics (TOG), 33(2),1-12.

569 Wang, Z., Li, Q., Sun, L., Li, S., Yang, Y., & Liu, S. (2015). The most typical shape of oceanic mesoscale eddies
570 from global satellite sea level observations. Frontiers of Earth Science, 9(2), 202-208. doi: 10.1007/s11707-014-
571 0478-z.

572 Woodring, J., Petersen, M., Schmeiber A., Patchett J., Ahrens J., Hagen H., (2016). In Situ Eddy Analysis in a High-
573 Resolution Ocean Climate Model, IEEE Transactions on Visualization & Computer Graphics, 22(1), 857-866,
574 doi:10.1109/TVCG.2015.2467411

575 Xiu, P., Chai, F., Shi, L., Xue, H., & Chao, Y. (2010). A census of eddy activities in the South China Sea during
576 1993–2007. Journal of Geophysical Research: Oceans, 115(C3). C03012, doi:10.1029/2009JC005657

577 Xu, C., Shang, X. D., & Huang, R. X. (2011). Estimate of eddy energy generation/dissipation rate in the world
578 ocean from altimetry data. Ocean Dynamics, 61(4), 525-541.

579 Yang, G., Wang, F., Li, Y., Lin, P., (2013). Mesoscale eddies in the northwestern subtropical Pacific Ocean:
580 Statistical characteristics and three-dimensional structures. Journal of Geophysical Research: Oceans, 118(4): 1906–
581 1923.

582 Yi, J., Du, Y., Zhou, C., Liang, F., & Yuan, M. (2015). Automatic Identification of Oceanic Multieddy Structures
583 From Satellite Altimeter Datasets. IEEE JSTARS, 8(4): 1555-1563.

584 Zhai, X., Johnson, H. L., & Marshall, D. P. (2010). Significant sink of ocean-eddy energy near western boundaries.
585 Nature Geoscience, 3(9), 608-612.

586 Zhang, C. H., Xi, X. L., Liu, S. T., et al. (2014). A mesoscale eddy detection method of specific intensity and scale
587 from SSH image in the South China Sea and the Northwest Pacific. Science China: Earth Sciences, 57: 1897–1906,
588 doi: 10.1007/s11430-014-4839-y.

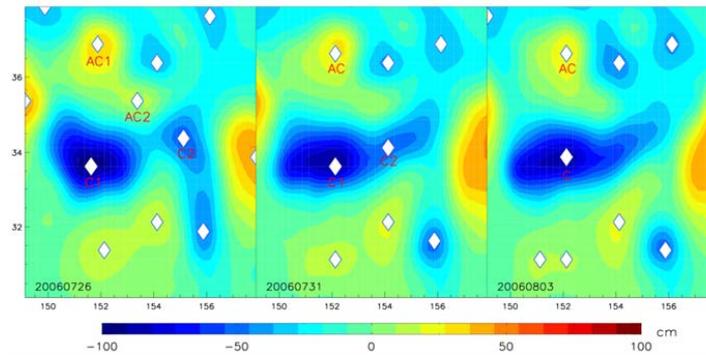
589 Zhang, Z., Zhang, Y., Wang, W., & Huang, R. X. (2013). Universal structure of mesoscale eddies in the ocean.
590 Geophysical Research Letters, 40(14), 3677-3681. doi: 10.1002/grl.50736.

591

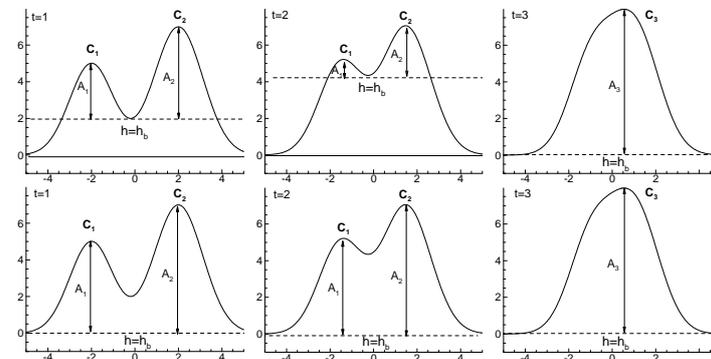
592

593   Table 1. The census of long-lived eddies, where "Amp" represents the amplitude threshold used in eddy detection;
594   and "C" and "AC," respectively, represent cyclonic and anticyclonic eddies.

Amp	AC (>100 d)	C (>100 d)	AC (>400 d)	C (>400 d)
1 cm	7290	3627	198	22
3 cm	7118	3550	194	21

595

596

597   Figure 1. The evolutions of amplitudes and areas of eddies from July 5 to August 3, 2006 (after Li et al. 2014). Two
598   anticyclonic eddies AC1 and AC2 merged into a single eddy on July 31, 2006. And, two cyclonic eddies C1 and C2
599   merged into a single one on August 3, 2006.

600

601

602   Figure 2. Top panels: Time evolution of two merging eddies revealed by the mononuclear eddy identification
603   without segmentation. Bottom panels: Time evolution of two merging eddies revealed by the mononuclear eddy
604   identification with segmentation.

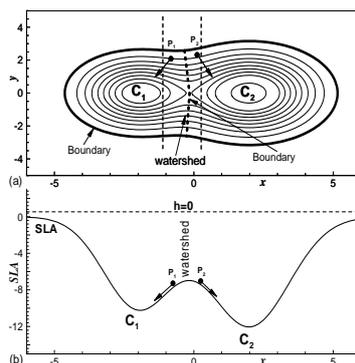



Figure 3. (a) Watershed as the natural division of eddies. (b) The particles on the watershed flow downward to the eddy centres. After Li and Sun (2015).

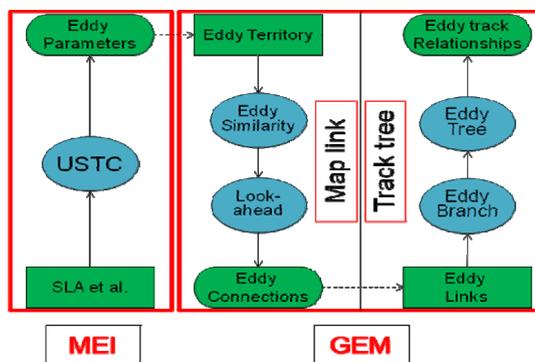



Figure 4. Flow chart of the systems. Mononuclear Eddy Identification (MEI) uses SLA to identify eddies via the Universal Splitting Technology for Circulations (USTC) method. The GEM, which has two independent parts of "Map link" and "Track tree," then uses the previously-identified eddies for tracking.

614

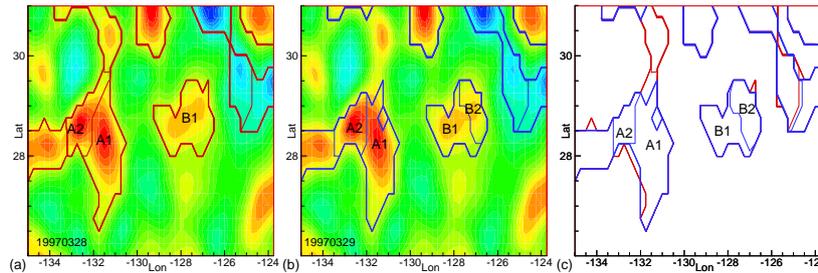
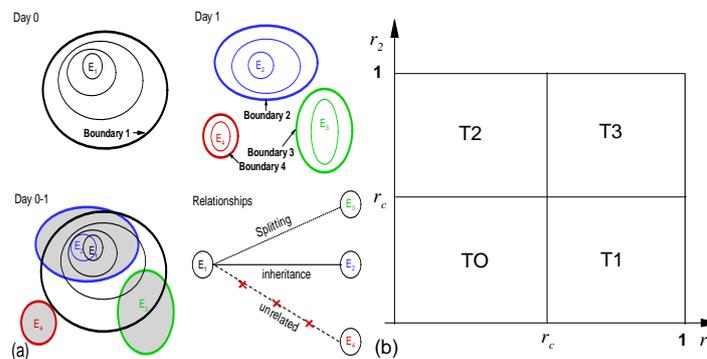



Figure 5. Sketch of eddy overlaps. (a) The SLA map (shading) and the boundary of eddies (red curves) on March 28, 1997. (b)The SLA map (shading) and the boundary of eddies (blue curves) on March 29, 1997. (c) The intersection of eddy territories by overlap eddy identification maps.

615
616
617

618

619

620   Figure 6. Sketch of eddy similarities. (a) The sketch of eddy overlaps. Eddy E_1 (black) is the eddy identified on day
621   0, where the thin contours represent the eddy parameter (e.g., the SLA value). The thick contour represents the eddy
622   boundary. Eddies E_2 (blue), E_3 (green) and E_4 (red) are identified on day 1. We consider the overlay between the two
623   eddies on different days to evaluate the similarity between them. (b) There are four types (T0-T3) according to the
624   values of r_1, r_2 and the critical value r_c.

625

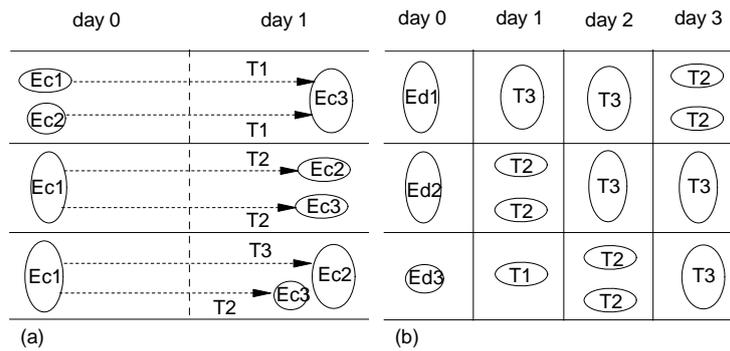



Figure 7. (a) Three typical cases of successors. (b) Different successors corresponding to different numbers of "look-ahead" days.

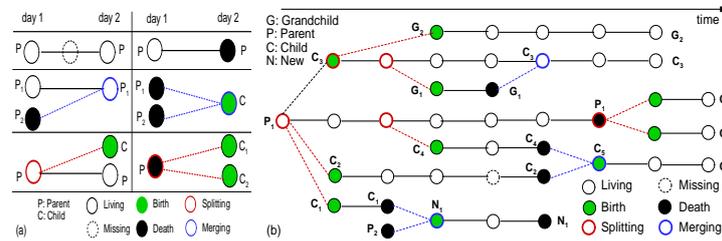
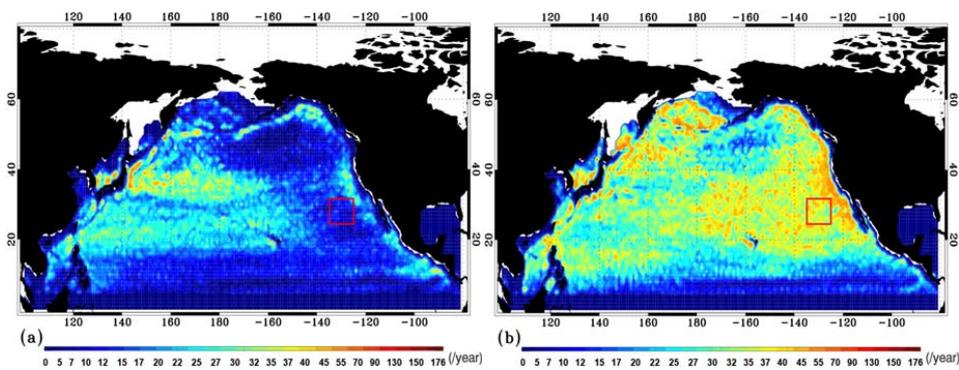
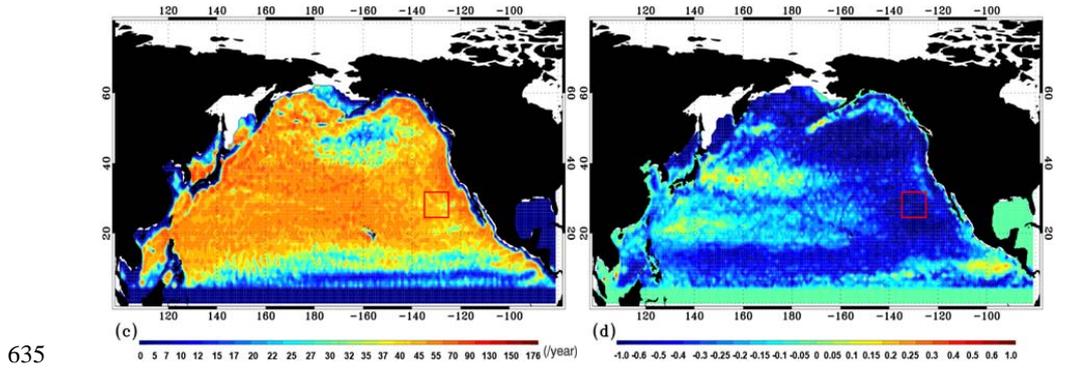
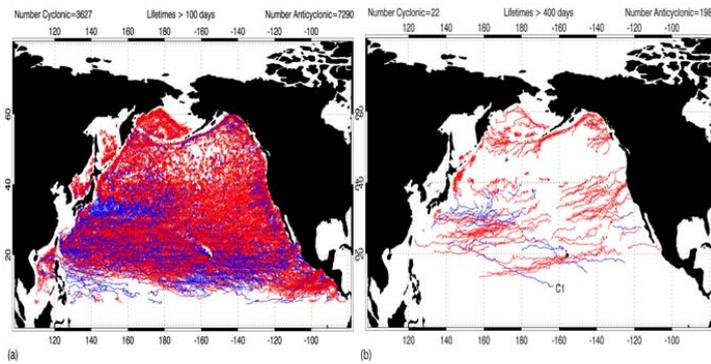
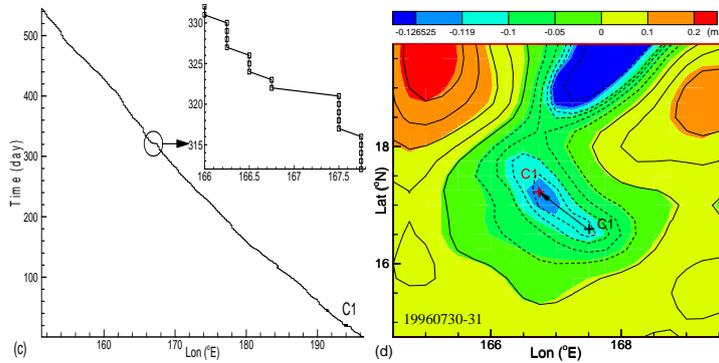



Figure 8. The logical genealogy of an ocean eddy with six states: birth, death, living, missing, splitting, and merging.
(a) The logical relationships of eddies between two days. (b) The logical genealogy evolution model of the eddies.

635

636   Figure 9 (a) The number of cyclonic eddy extrema on each 1°×1° grid per year. (b) Same as (a), except for

637   anticyclonic eddies. (c) Same as (a), except for the total number of eddies. (d) The ratios of difference in number of

638   cyclonic and anticyclonic eddies to the total eddies (A logarithmic scale is used).

639

640

641

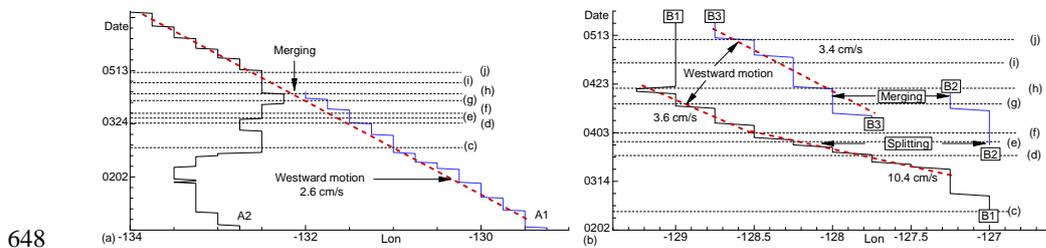
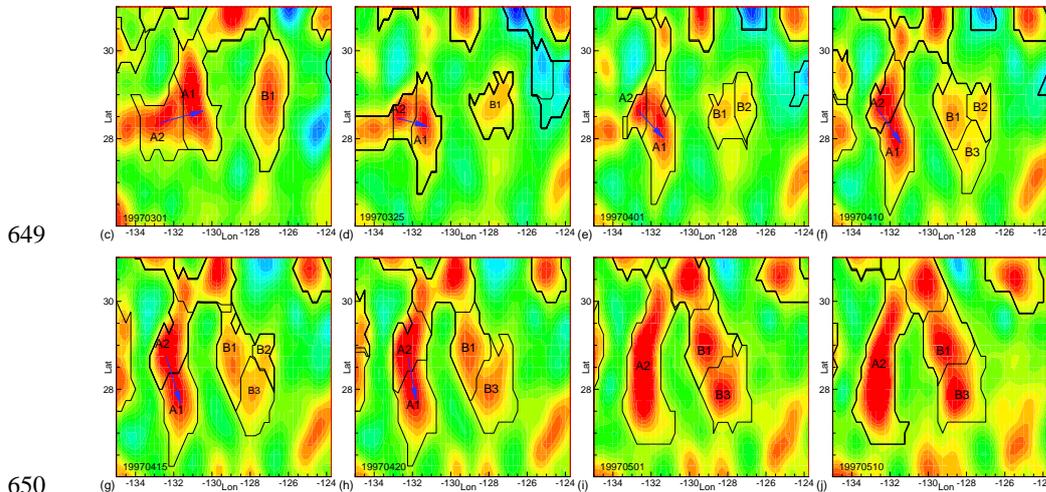
642   Figure 10. (a) Tracks of long-lived (>100 days) eddies. (b) Tracks of long-lived (>400 days) eddies. In (a) and (b),

643   blue color marks cyclonic eddies, and red color marks anticyclonic eddies. (c) The track of eddy C1. Note the

644   sudden jump from 167.5°E to 166.75°E on July 31, 1996. (d) The SLA fields on July 30 (contours) to 31 (shading),

645   using the same intervals for the contours and the shadings. The eddy centers are marked by a black cross (July 30)

646   and a red cross (July 31).

647

650

651   Figure 11. The dynamic evolutions of two groups of eddies, which are located in the red boxes in Fig. 9. (a) Two

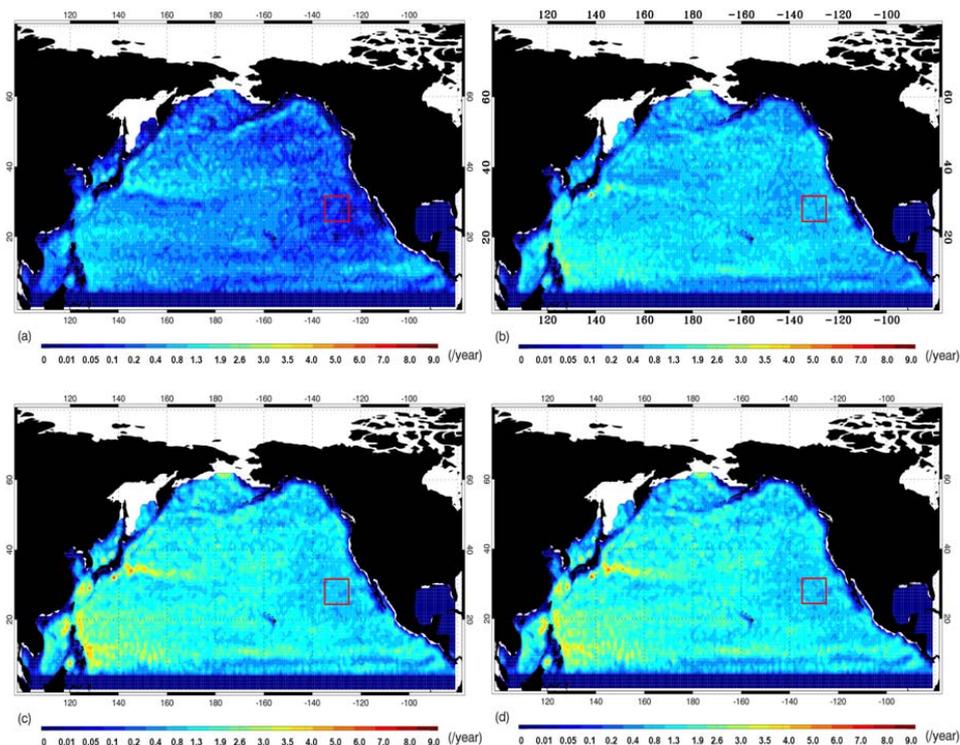
652   eddies, A1 and A2, approached each other, and A1 merged with eddy A2, where the blue arrows indicate that the

653   eddy centers rotated clockwise during the merging process. (b) In the mean time, eddy B1 split into two small eddies.

654   (c)-(j) The evolutions of SLA fields and eddies. Note that eddies A1 and A2 had clockwise rotations when they

655   approached each other, as indicated by the blue arrows in (c)-(h).

656

657

658

659 Figure 12. The frequencies of dynamic processes per $1^{\circ} \times 1^{\circ}$ grid element. (a) The merging frequency for cyclonic

660 eddies. (b) The merging frequency for anticyclonic eddies. (c) The merging frequency for all eddies. (d) The

661 splitting frequency for all eddies.

662

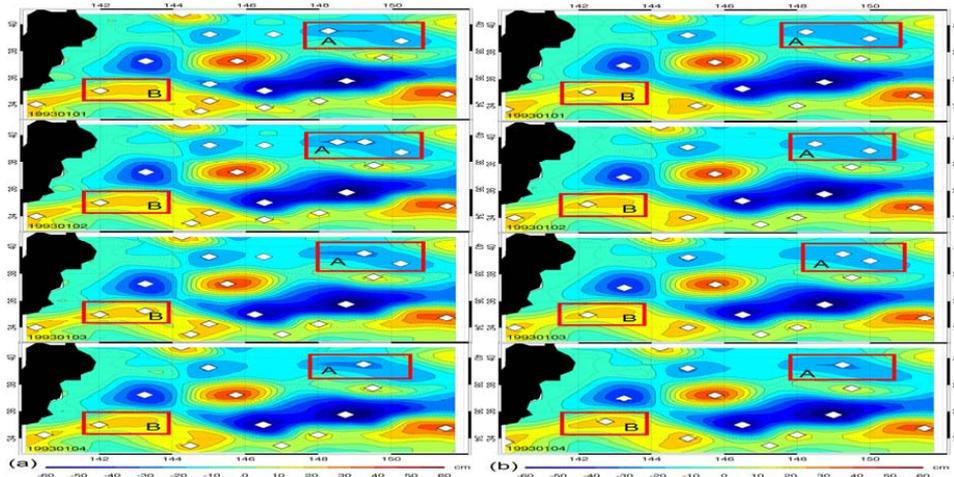



Figure 13. Comparison of the non-smoothed (a) and smoothed SLA data (b) from January 1 to January 4, 1993.
Note that small noises affected the eddy detection.

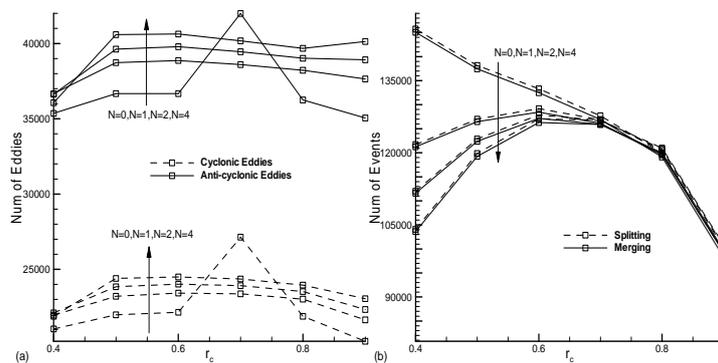



Figure 14. (a) Number of eddies (lifetime > 30 days) vs. the critical value r_c and look-ahead time N. (b) Number of
merging and splitting events vs. the critical value r_c and look-ahead time N.