INTRODUCTION

The marine geoid (\(N\)) is an equipotential surface of the earth’s gravity field to represent the mean sea level without the influence of tides, currents and winds. With a hypothetical situation that if the sea surface elevation was governed only by geostrophy and coincided one particular equipotential taken as the geoid, then there would be no flow on the geoid. This hypothesis has never been tested since the marine geoid is generally considered as an imaginary surface by the oceanographic community. After satellites coming into practice, the surface \(N\) has been determined by various gravity models using satellite observations such as Gravity field and steady-state Ocean Circulation Explorer (GOCE) (Johannessen et al., 2003), Gravity Recovery and Climate Experiment (GRACE) (Jet Propulsion Laboratory, 1998) and combined GOCE-GRACE (Bruinsma et al., 2013). Thus, the hypothesis can be tested through calculating the absolute geostrophic current (AGC) on \(N\) from temperature (\(T\)) and salinity (\(S\)) fields using existing inverse methods such as the \(\beta\)-spiral (Stommel & Schott, 1977), box model (Wunsch, 1978), minimum energy (Chu, 2018) and P-vector (Chu, 1995; Chu et al., 1998a; Chu & Li, 2000; Yuan et al., 2014).

A new data set ‘World ocean annual mean absolute geostrophic velocity on marine geoid of EIGEN-6C4 from WOA13’ (https://doi.org/10.14284/398) (Chu, 2020) was generated from two open data sets: (a) ‘World ocean geostrophic velocity inverted from World Ocean Atlas 2013 with the P-vector method’ (NCEI accession 0,121,576). With the given non-positive values of \(N\) in the oceans, absolute geostrophic currents (\(u\), \(v\)) are easily obtained on \(N\) with 1°×1° resolution from the second data set except the equatorial zone (5oS – 5oN) due to the non-existence of the geostrophic balance. Altogether, the data set contains 15,868 (\(u\), \(v\)) data pairs. The data set shows that the hypothetical situation that there would be no flow on the geoid does not exist.

KEYWORDS
absolute geostrophic velocity, EIGEN-6C4 gravity model, marine geoid, P-vector inverse method, WOA13

1 | INTRODUCTION

This data set was established from two open sources: (a) geoid undulation (\(N\)) from EIGEN-6C4, which is a static global combined gravity field model up to degree and order 2,190, and (b) ‘World ocean geostrophic velocity inverted from World Ocean Atlas 2013 with the P-vector method’ (NCEI accession 0,121,576). With the given non-positive values of \(N\) in the oceans, absolute geostrophic currents (\(u\), \(v\)) are easily obtained on \(N\) with 1°×1° resolution from the second data set except the equatorial zone (5oS – 5oN) due to the non-existence of the geostrophic balance. Altogether, the data set contains 15,868 (\(u\), \(v\)) data pairs. The data set shows that the hypothetical situation that there would be no flow on the geoid does not exist.
DATA PRODUCTION METHODS

2.1 EIGEN-6C4

The EIGEN-6C4 model ( Förste et al., 2014; Ince et al., 2019) is on the website http://icgem.gfz-potsdam.de/home to produce global static geoid undulation (N) data set. The geoid (N) is given by the Bruns’ formula (Vaniček & Krakiwsky, 1986).

\[ N(\lambda, \varphi) = \frac{T(\lambda, \varphi, 0)}{g}, \quad g = 9.81 \text{ m s}^{-2} \quad (1) \]

where (\lambda, \varphi) are longitude and latitude; T is the static disturbing gravity potential and given by Kostelecký et al., (2015)

\[ T(r, \lambda, \varphi) = \frac{GM}{r} \sum_{l=2}^{\infty} \sum_{m=0}^{l} \left( \frac{R}{r} \right)^{l} \left[ \left( C_{l,m} - C_{l,m}^{c} \right) \cos m\lambda + S_{l,m} \sin m\lambda \right] P_{l,m}(\sin \varphi) \quad (2) \]

where \( G = 6.674 \times 10^{11} \text{ m}^3 \text{kg}^{-1} \text{s}^{-2} \) is the gravitational constant; \( M = 5.9736 \times 10^{24} \text{ kg} \) is the mass of the earth; \( R = 6.3781364 \times 10^6 \text{ m} \) is the earth radius; \( r \) is the radial distance with \( z = r - R \); \( P_{l,m}(\sin \varphi) \) is the Legendre associated functions with \( l, m \) the degree and order of the harmonic expansion; \( (C_{l,m}, C_{l,m}^{c}, S_{l,m}) \) are the harmonic geopotential coefficients (Stokes parameters) with \( C_{l,m}^{c} \) belonging to the reference ellipsoid (Stokes, 1849). The EIGEN-6C4 model was developed jointly by the GFZ Potsdam and GRGS Toulouse up to degree and order 2,190. Following the instruction in the website, the author ran the EIGEN-6C4 model in 1°x1° resolution for 17 s to get the global N (Figure 1).

2.2 P-vector inverse method

The P-vector inverse method was proposed in 1995 and described in detail in a book (Chu, 2006). Readers may ignore this section if there is no intention to use this method to calculate AGC from (T, S) data. This method is outlined as follows (Chu, 1995, 2006; Chu & Fan, 2015a). Let the earth-fixed local coordinates (x, y, z) be used in the P-vector calculation with x-axis in the zonal direction (eastward positive), y-axis in the meridional direction (northward positive), and z-axis in the vertical (upward positive). Let \( \mathbf{V} \) be the velocity vector and \( \rho \) the density. Large-scale motions in the ocean interior are geostrophic and hydrostatic balanced, and density conserved,

\[ \mathbf{V} \cdot \nabla \rho = 0. \quad (3) \]
These conditions lead to the conservation of the potential vorticity \( q = f \rho / \partial z \),

\[
V \cdot \nabla q = 0
\]  

(4)

where \( f \) is the Coriolis parameter and \( V \) is the three-dimensional gradient operator. Equations (3) and (4) show that the vector \( V \) is perpendicular to both \( \nabla \rho \) and \( \nabla q \) and therefore parallel to \( \nabla \rho \times \nabla q \). Thus, the velocity vector \( V \) is given by

\[
\mathbf{V} = \gamma(v, y, z) \mathbf{P}, \quad \mathbf{P} = \frac{\nabla \rho \times \nabla q}{|\nabla \rho \times \nabla q|}, \quad \text{(5)}
\]

where the unit vector \( \mathbf{P} \) is calculated from hydrographic data; \( \gamma \) is the speed parameter with \( |\gamma| \) the speed. After the unit vector \( \mathbf{P} \) is determined, the speed parameters \( \gamma(k) = \gamma(x, y, z_k) \) and \( \gamma(m) = \gamma(x, y, z_m) \) satisfy the algebraic equations.

\[
\gamma(k)p_x^k - \gamma(m)p_x^m = \Delta u_{k, m}, \quad \text{(6a)}
\]

\[
\gamma(k)p_y^k - \gamma(m)p_y^m = \Delta v_{k, m}, \quad \text{(6b)}
\]
due to the thermal wind relation between \( z_k \) and \( z_m \).

\[
\Delta u_{k, m} \equiv \frac{g}{f \rho_0} \int_{z_m}^{z_k} \frac{\partial \rho}{\partial y} \, dz', \quad \text{(7a)}
\]

\[
\Delta v_{k, m} \equiv -\frac{g}{f \rho_0} \int_{z_m}^{z_k} \frac{\partial \rho}{\partial x} \, dz'. \quad \text{(7b)}
\]

If the determinant of the two linear algebraic equations (6a) and (6b) is non-zero,

\[
\begin{vmatrix}
  p_x^k & p_x^m \\
  p_y^k & p_y^m
\end{vmatrix} \neq 0, \quad \text{(8)}
\]

that is the \( P \)-vector spiral exists (Chu, 2000), the speed parameter \( \gamma \) at these two levels \( \gamma(k) \) and \( \gamma(m) \) can be determined after solving the linear algebraic equations (6a) and (6b), and in turn the horizontal velocity (i.e. AGC).

The traditional AGC calculation using the \( P \)-vector method employs the least squares fitting over the entire water column. This could introduce errors into the \( P \)-vector AGCs, because the potential density and potential vorticity are not conserved in the surface mixed layer, because the whole-column fitting could cause the fitting to be biased towards the large vectors. To test the accuracy, a second computation for AGC (\( u^*, v^* \)) was conducted with the \( P \)-vector calculation in the subsurface water column from \( z = -500 \) m to the bottom and calculate the upper ocean \((z > -500 \) m) geostrophic currents using the dynamic height calculation. For each \( z \) level, the relative root mean square difference (RRMSD) is calculated to represent the accuracy. Here, \( N_k \) is the number of the grid points for the \( z_k \)-level; \( N \) is the total number of the grid points for the three-dimensional ocean. Figure 2 shows that the RRMSD reduces rapidly with depth from 0.12 near the sea surface to 0.013 at \( z = -5,500 \) m. Besides, the number of observational data to build WOA13 (T, S) climatology is relatively low in high latitudes. This may cause large error bars there. Using the \( P \)-vector method, a global three-dimensional AGC equations (9)}

\[
\begin{vmatrix}
  p_x^k & p_x^m \\
  p_y^k & p_y^m
\end{vmatrix} \neq 0, \quad \text{(8)}
\]
data set was established from WOA13 and published on the NCEI website (https://data.nodc.noaa.gov/cgi-bin/iso?id=gov.noaa.nodc:0121576) for public use (Chu & Fan, 2015a,b).

2.3 Identification of AGC on N

The aforementioned P-vector method is simple and effective to compute AGC from \((T, S)\) fields. If we use the existing inverted AGC field from the open data set on the NCEI website (https://data.nodc.noaa.gov/cgi-bin/iso?id=gov.noaa.nodc:0121576), the annual mean AGC data set \((WOZV13\_annual.nc)\) should be downloaded since the static gravity model EIGEN-6C4 provides the climatological annual mean \(N\) field. This open AGC data set is in the \(z\)-coordinate system from \(z = 0\) (top) to \(z = -8,900\) m (bottom). Therefore, the AGC on \(N\) in the oceans can be obtained from this data set with only non-positive \(N\) values. Because positive \(N\) is out of the range of \(z \geq 0 \geq -8,900\) m, no data would be available from the climatological AGC field.

3 DATA RECORDS

This global data set for the AGC on \(N\) is publicly available at the Flanders Marine Institute website (https://doi.org/10.14284/398) as text and NetCDF files, which include data citation, data set identifiers, metadata and ordering instructions. The 15,868 AGC vectors \((u_N, v_N)\) (Figure 3), histograms of \((u_N, v_N)\) (Figure 4) and associate statistical parameters (Table 1) clearly show that the water moves on \(N\).

The gappy areas in Figure 3 include the area with \((N > 0\), i.e. above the sea surface\) where no data exist in the annual mean AGC data set \((WOZV13\_annual.nc)\) and the equatorial area \((5^\circ S - 5^\circ N)\) where the geostrophic balance does not exist. A possible approach to determine AGC on the marine geoid for \(N > 0\) is to let the \(N\) surface and the \(S\) surface not relative to the same ellipsoid, that is to submerge the \(N\) surface into the ocean and to calculate the AGCs on the entire \(N\) surface. This will lead to a classical problem in physical oceanography about the reference level, which is beyond the scope of this paper.

4 TECHNICAL VALIDATION

The two models (EIGEN-6C4 and P-vector) are well established for many years. The EIGEN-6C4 is a geodetic community model after validation (Fürste et al., 2014; Ince et al., 2019; Klokočník et al., 2020; Kostelecký et al., 2015; Pal et al., 2016; Steffen et al., 2017). The P-vector method was evaluated (Chu et al., 1998b) and applied to calculate the absolute velocity from hydrographic data for the South China Sea (Chu et al., 1998a; Chu & Li, 2000), the Japan

FIGURE 3 Velocity vectors of AGC at the geoid undulation \((N)\) obtained from the open data set (from Figure 2b in Chu 2020b)
In conjunction with the wind forcing, the P-vector method is also used to calculate the global volume transport (Chu & Fan, 2007). To reduce error, a variational P-vector method was developed (Chu et al., 2001b).

The climatological annual mean AGC on the marine geoid \((u_N, v_N)\) along with the EIGEN-6C4 geoid \((N)\) data (unit: m) on 1°×1° resolution was published by the Flanders Marine Institute at the website https://doi.org/10.14284/398 for public use. Each data set has text and netCDF formats with \((uvN.nc, uvN.txt)\) for the annual mean AGC on \(N\) and \((EIGEN-6C4.nc, EIGEN-6C4.txt)\) for the geoid data. Readers are also encouraged to calculate the monthly mean AGC on the marine geoid from the two open sources, http://icgem.gfz-potsdam.de/home and https://data.nodc.noaa.gov/cgi-bin/iso?id=gov.noaa.nodc:0121576 with two MATLAB codes. The first code pv.m is used to calculate the AGC from \((T, S)\) fields (Appendix 1). The second code getuvN.m is used to get AGC on the marine geoid undulation \(N\) [i.e. \((u_N, v_N)\)].
Acknowledgements
The author thanks the International Centre for Global Earth Models (ICGEM) for the EIGEN-6C4 geoid undulation data, and the NOAA National Centers for Environmental Information (NCEI) for the World Ocean Atlas 2013 temperature and salinity fields and corresponding absolute geostrophic velocity inverted by the P-vector method. He also thanks Chenwu Fan at the Naval Postgraduate School for computational assistance, and Carolien Knockaert at the Flanders Marine Institute for quality control and publication of this data set on the Flanders Marine Institute website. The Research Office of the Naval Postgraduate School supports the publication in the Geoscience Data Journal.

OPEN RESEARCH BADGES
This article has earned an Open Data badge for making publicly available the digitally shareable data necessary to reproduce the reported results. The data are available at https://doi.org/10.14284/398.

ORCID
Peter C. Chu  https://orcid.org/0000-0002-3101-3555

REFERENCES
Bruinsma, S.L., Förste, C., Abrikosov, O., Lemoine, J.M., Marty, J.C., Mulet, S. et al. (2013) The new ESA satellite-only gravity field model via the direct approach. Geophysical Research Letters, 40, 4607–4612.
Chu, P.C. (1995) P-vector method for determining absolute velocity from hydrographic data. Marine Technical Society Journal, 29(3), 3-14.
Chu, P.C. (2000) P-vector spirals and determination of absolute velocities. Journal of Oceanography, 56, 591–599.
Chu, P.C. (2006) P-vector inverse method. Heidelberg, Berlin: Springer Press, 605 pp.
Chu, P.C. (2018) Determination of dynamic ocean topography using the minimum energy state. Universal Journal of Geosciences, 6(2), 25–39. https://doi.org/10.13189/ujg.2018.060201
Chu, P.C. (2020) World ocean annual mean absolute geostrophic velocity on marine geoid of EIGEN-6C4 from WOA13. Flanders Marine Institute. https://doi.org/10.14284/398
Chu, P.C. & Fan, C.W. (2007) An inverse model for calculation of global volume transport from wind and hydrographic data. Journal of Marine Systems, 65, 376–399.
Chu, P.C. & Fan, C.W. (2015a) Absolute geostrophic velocity inverted from World Ocean Atlas 2013 (WOAV13) with the P-vector method. Geoscience Data Journal, 2, 78–82.
Chu, P.C. & Fan, C.W. (2015b) World ocean geostrophic velocity inverted from world ocean atlas 2013 with the P-vector method. NOAA National Centers for Environmental Information. https://data.nodc.noaa.gov/cgi-bin/iso?id=gov.noaa.nodc:0121576
Chu, P.C., Fan, C.W. & Cui, W.J. (1998b) Evaluation of P-vector method using modular ocean model (MOM). Journal of Oceanography, 54, 185–198.
Chu, P.C., Fan, C.W., Lozano, C.J. & Kerling, J. (1998a) An AXBT survey of the South China Sea. Journal of Geophysical Research, 103, 21637–21652.
Yuan, D.L., Zhang, Z.C., Chu, P.C. & Dewar, W.K. (2014) Geostrophic circulation in the tropical north Pacific Ocean based on Argo profiles. *Journal of Physical Oceanography*, 44, 558–574.

Zweng, M.M., Reagan, J.R., Antonov, J.I., Locarnini, R.A., Mishonov, A.V., Boyer, T.P. et al. (2013). *World Ocean Atlas 2013, Volume 2: Salinity*. Levitus S. Ed.; Mishonov A, Technical Ed.; NOAA Atlas NESDIS 74, 39 pp.

**How to cite this article:** Chu PC. World ocean annual mean absolute geostrophic velocity on marine geoid of EIGEN-6C4 from WOA13. *Geosci Data J*. 2022;9:131–139. [https://doi.org/10.1002/gdj3.124](https://doi.org/10.1002/gdj3.124)
APPENDIX 1

Matlab Code (getuvN.m)

```matlab
close all;

ncid=netcdf.open('EIGEN-6C4.nc','nowrite');

vid=netcdf.inqVarID(ncid,'GEOID');
N=netcdf.getVar(ncid,vid);

ncid=netcdf.open('WOZV13_annual.nc','nowrite');

vid=netcdf.inqVarID(ncid,'lat');
lat1=netcdf.getVar(ncid,vid);

vid=netcdf.inqVarID(ncid,'lon');
lon1=netcdf.getVar(ncid,vid);

vid=netcdf.inqVarID(ncid,'u');
u3d=netcdf.getVar(ncid,vid);

vid=netcdf.inqVarID(ncid,'v');
v3d=netcdf.getVar(ncid,vid);

vid=netcdf.inqVarID(ncid,'depth');
dep=netcdf.getVar(ncid,vid);

netcdf.close(ncid);

u3d(u3d>1e20)=NaN; v3d(v3d>1e20)=NaN;

u3d=permute(u3d,[3,2,1]);
v3d=permute(v3d,[3,2,1]);

ii=find(lon1==20);
lon1=cat(1,lon1(ii:end),lon1(1:ii-1)+360);
u3d=cat(3,u3d(:,:,ii:end),u3d(:,:,1:ii-1));
v3d=cat(3,v3d(:,:,ii:end),v3d(:,:,1:ii-1));
[mz,my,mx]=size(u3d);

u1=NaN*ones(my,mx); v1=u1;

for j=1:my
    for i=1:mx
        if N(j,i)<=0
            uN(j,i)=interp1(dep,u3d(:,j,i),N(j,i));
            vN(j,i)=interp1(dep,v3d(:,j,i),N(j,i));
        end
    end
end

save uvN.mat lon1 lat1 uN vN;
```
function [u,v,w,invmask,ug,vg]=pv(lat,lon,temp3d,salt3d,depth,topo);

global mx my mz ff ktop sims small minlat small_rhotz power fac_scale dx3d dy dz3d rho0;
[mz,my,mx]=size(temp3d);
if(~exist('depth'))
    depth=[0,10,20,30,50,75,100,125,150,200,250,300;100:1500,1750,2000;50:5500]';
else
    depth=abs(depth);
end
mx1=mx-1; my1=my-1; mz1=mz-1;
%-------------------------------------------------------------
ktop=5; rho0=1028; small=1e-7; minlat=5; sims=1e-7;
fac_scale=1e-6; power=2; small_rhotz=1.0e-20;
d2r=pi/180; omega = 7.292e-5; radius = 6.371e6; pref=0;

gridsize=lat(2)-lat(1);
if(lat(2)>lat(1))
ydir=1;
else
    ydir=-1; gridsize=-gridsize;
end

ff=2*omega*sin(lat*d2r);
j1=find(lat>=0 & lat<minlat); j2=find(lat<0 & lat>-minlat);
if(~isempty(j1))
    ff(j1)=2*omega*sin(minlat*d2r);
end
if(~isempty(j2))
    ff(j2)=-2*omega*sin(minlat*d2r);
end
ddep=diff(depth);
[lt,dz2d]=meshgrid(lat,ddep);
dz3d=zeros(mz1,my,mx); dx3d=zeros(mz,my,mx1);
for i=1:mx
    dz3d(:,:,i)=dz2d;
end
dy=(lat(2)-lat(1))*d2r*radius;
dx=(lon(2)-lon(1))*d2r*radius*cos(lat*d2r);
for j=1:my
    dx3d(:,:,j)=dx(j)*ones(mz,1,mx1);
end