Lee wave generation rates in the deep ocean
Corwin J. Wright\textsuperscript{1,2}, Robert B. Scott\textsuperscript{1,3,4}, Pierre Ailliot\textsuperscript{5} and Darran Furnival\textsuperscript{1}

Corresponding author: C. J. Wright, Centre for Space, Atmospheric and Oceanic Science, University of Bath, Bath, UK (corwin.wright@trinity.oxon.org)

\textsuperscript{1}Laboratoire de Physique des Océans, Université de Bretagne Occidentale, Brest, France
\textsuperscript{2}now at Centre for Space, Atmospheric and Oceanic Science, University of Bath, Bath, UK
\textsuperscript{3}CNRS, France
\textsuperscript{4}Institute for Geophysics, The University of Texas at Austin, Austin, Texas, USA
\textsuperscript{5}Laboratoire de Mathématiques, Université de Bretagne Occidentale, Brest, France
Using the world’s largest dataset of in-situ ocean current measurements, combined with a high resolution topography roughness dataset, we use a model-assisted hierarchical clustering methodology to estimate the global lee wave generation rate at the ocean floor. Our analysis suggests that internal wave generation contributes $0.75 \pm 0.19$ TW ($\pm 2$ st. dev.) to the oceanic energy budget, but with a strong dependence on the Brunt-Väisälä (buoyancy) frequency climatology used. This estimate is higher than previous calculations, and suggests that internal wave generation may be a much more significant contributor to the global oceanic mechanical energy budget than had previously been assumed. Our results imply that lee wave generation and propagation may be a dominant sink of as at least half and potentially the overwhelming majority of ocean surface wind work on the geostrophic circulation.
1. Introduction

The global mechanical energy budget is a key component in our understanding of the ocean general circulation, yet remains to a large degree poorly quantified [Munk and Wunsch, 1998; Wunsch and Ferrari, 2004; Ferrari and Wunsch, 2009]. It has been shown that wind work on the surface general circulation provides a dominant power source [Xu and Scott, 2008; von Storch et al., 2007; Scott, 1999] of around 0.75 – 1 TW [Scott and Xu, 2009; Hughes and Wilson, 2008; Wunsch, 1998], but isolating the mechanism by which the wind power input is dissipated at sub-surface depths is a much greater challenge. The precise dissipation mechanism has important ramifications: if deposited in the right place, this power could provide an energy source for the turbulent diapycnal mixing required to drive the meridional oceanic overturning circulation [Kuhlbrodt et al., 2007; Saenko et al., 2012].

There are several candidates for the dissipation mechanism. These include ageostrophic instabilities in the ocean interior [Müller et al., 2005], nonlinear coupling to internal gravity waves [Bühler and McIntyre, 2005], turbulent bottom boundary layer generation by mesoscale currents [Wright et al., 2012, 2013], lee wave generation by the flow of mesoscale currents over rough topography at the ocean floor [Scott et al., 2011], or some combination of these and other processes.

We here focus on the lattermost such process. The generation of lee waves at the ocean floor is a known sink of current energy in the deep ocean [Scott et al., 2011; Nikurashin et al., 2012]. Mesoscale currents flow over the ocean bed, generating waves in the lee of rough topography which propagate away from the source [Bell, 1975]; this transports en-
ergy and momentum and provides a mechanism for the transfer of energy from mesoscale
flows to turbulent lengthscales. However, to date the rate at which this energy is trans-
ferred from the mean flow to lee waves has not been assessed on a global scale with in-situ
data. Estimates using ocean current speeds from the HYCOM model [Chassignet et al.,
2007] suggest a value $\sim 0.4 \pm 0.1$ TW [Scott et al., 2011], where the error bars represent
sensitivity to choice of roughness dataset, year of currents, and various other parameters,
whilst those from a model developed at the GFDL suggest $\sim 0.2$ TW [Nikurashin and Fer-
rari, 2011]. We believe both of these studies underestimate the total lee wave generation
power, since global circulation models tend to exhibit significant negative biases in bottom
current speeds [Scott et al., 2010, 2011]. This may at least partially be due to the tendency
of data-assimilative models to be driven and tuned primarily at near-surface levels, due
to the greater availability of observational data such as, e.g., satellite measurements and
ARGO floats at these depths.

Since the lee wave generation rate is critically dependent upon water velocity, this
probable underestimation is of great importance. Any increase in the estimate for this
quantity would suggest that lee wave generation is the dominant mechanism for energy
dissipation at the ocean floor. Consequently, we here compute the global lee wave energy
dissipation rate $G$ using data from the world’s largest collection of ocean current meter
time series, the Global Multi-Archive Current Meter Database (GMACMD) [Sen et al.,
2008; Wright et al., 2012, 2013]. Our calculation uses the HYCOM model for assistance
in extrapolating the lee wave generation rate beyond the geographic locations of the
current meters; this assistance is necessary due to the extreme spatial inhomogeneity
of the distribution of available measurements, with significant gaps particularly in the Southern and Pacific oceans.

2. Method

2.1. Pointwise Lee Wave Generations

Lee wave generation estimates are calculated using the linearised solutions of the density and vertical momentum equations to compute the work rate of topographic form drag \cite{Scott et al., 2011; Gill, 1982}. We use a small-scale topographic roughness dataset, of the 2D topographic power spectrum $P(k,l)$ \cite{Goff, 2010}, where $k$ and $l$ are the zonal and meridional wavenumbers of the topography respectively. We then integrate over the internal gravity waveband $|f_0| \leq \omega \leq N$, where $f_0$ is the Coriolis frequency, $N$ the Brunt-Väisälä (buoyancy) frequency, and $\omega = uk + vl$ the frequency of the generated wave, according to

$$G_i(t) = \rho_0 \int \int \frac{\omega}{\sqrt{k^2 + l^2}} P(k,l) \sqrt{N^2 - \omega^2} \sqrt{\omega^2 - f_0^2} dkdl,$$

where $G_i(t)$ is the generation rate for that current meter (in Watts) and $\rho_0 = 1035 \text{ kg m}^{-3}$ the density of water. Finally, we time-average $G_i(t)$ to obtain the time-mean generation rate for each meter $\bar{G}_i$. The horizontal current velocity $\mathbf{v} = (u, v)$ is derived from current meters or HYCOM as appropriate. Larger topographic features are expected to block the flow, reducing generation \cite{Wunsch, 1976}, and accordingly a correction factor is applied to compensate for this \cite{Welch et al., 2001}. Topographic data are not defined for the 35% of the ocean floor area which is considered to be too smooth to contribute significantly to the lee wave generation rate \cite{Scott et al., 2011}; consequently, we omit these areas from our calculation.
HYCOM current speeds are derived from five-daily-snapshots means of daily data-assimilative runs of the model on a nominal 1/12° Mercator grid in the horizontal and 32 hybrid layers in the vertical; Wright et al. [2012] contains further details relating to coordinates and forcings.

2.2. Brunt-Väisälä Frequency Data

Our primary calculation uses values of $N$ determined from the WOA2009 seasonal temperature and salinity climatology using equation 3.71 of [Gill, 1982] and averaged over the seasons, omitting negative values. We also include a secondary calculation in the article supplementary material, derived from estimates of $N$ computed using OCean Comprehensible Atlas (OCCA) data averaged over the period 2004-2006 [Forget, 2010]. The WOA2009-derived result is used as our primary estimate for consistency with Scott et al. [2011], Wright et al. [2012] and Wright et al. [2013]. Due to the slightly different spatial coverage of the WOA2009 and OCCA $N$ datasets, the two analyses required some standardisation to give equivalent results and remain consistent with the above-mentioned studies. To achieve this, gridpoints with values only in the OCCA climatology were removed, whilst gridpoints with values only in the WOA2009 climatology were duplicated in both analyses. These duplicated gridpoints make up around 8% of those used in the OCCA analysis.

2.3. Hierarchical Cluster Extrapolation

To extrapolate our individual current meter results to the whole ocean, we use the model-assisted hierarchical clustering methodology described by Wright et al. [2013] to generate a set of regions such that each contains at least one current meter.
First, a weighted interpoint distance matrix is computed using both the geographical distance and the difference in $G_i$ between each current meter pair in our dataset. A hierarchical tree ['dendrogram', Hastie et al., 2009] is then generated from this interpoint distance matrix. The tree is divided into a varying number of clusters; these clusters are defined using Ward’s linkage, which minimises the variance within each cluster to produce a set of geographically compact clusters which exhibit a similar $G_i$. We then compute the geographic centre of this cluster.

A separate estimate of $G_i$ is then computed for each point on a 1/3 degree Mercator grid using bottom current flow speeds from HYCOM [Chassignet et al., 2007]. Using this estimate, each grid point is assigned to the nearest cluster centre in a weighted space. The weighting is defined such that a change in $G_i$ of $1 \times 10^{-4}$ Wm$^{-2}$ is equivalent to a physical distance of 100 km, and was chosen empirically by examining sample maps of cluster output. The geographic centre derived from the current meters is used for the distance component of the calculation, whilst the change in $G_i$ is determined using the HYCOM-derived $G_i$ at this centre point.

This method produces a set of regions defined to each contain at least one current meter. For each such region, we compute the mean generation rate for the meters in this region $G_r$, and scale it by the area of the corresponding region; the sum of these area weighted means provides a first estimate of the globally integrated $G$ value. However, we expect there to be some bias in the geographical locations of our physical measurements [Sen et al., 2008; Wright et al., 2013; Holloway et al., 2011]. In order to estimate this bias, we compute an analogous estimate of $G$ using HYCOM data at the current meter locations and compare
this to a global HYCOM-derived estimate of $G$. This gives a correction factor, which we
assume to be a reliable estimate of the bias of the methodology and which is applied to
the first estimate of globally integrated $G$ computed from current meter data. The value
resulting from this correction is our final estimate $G$ of the globally integrated value. This
bias correction factor generalises that applied by Arbic et al. [2009], and reduces to it in
the case of a single cluster including data covering the whole World Ocean.

A maximum-weighted-distance criterion of 5000 km is imposed such that points deemed
too distant from any current meter record are not considered; HYCOM estimates suggest
these omitted regions should contribute no more than 3–5% of global lee wave generation.

2.4. Uncertainty Estimates

Uncertainty estimates are based upon statistical bootstrapping of the data [Efron, 1979].
This technique uses repeated random sampling, with replacement, from the datasets to
generate artificial $G$ values, allowing us to obtain an empirical distribution for our esti-
mate. Synthetic current meter values are generated based upon HYCOM values, using
a conditional bootstrap based upon nearest-neighbour proximity. The method generates
artificial values for our global estimate of $G$; the resulting empirical distribution of these
values is used to compute the confidence intervals shown on figure 3. These consequently
reflect the uncertainty due to the observed differences between HYCOM and current meter
values.

3. Methodology Assessment

We first wish to assess whether this model-assisted calculation is plausible. For it to
be a suitable method, the absolute model velocity estimates are not necessarily impor-
tant, but the spatial distribution and relative values of current meter measurements and
model estimates must be broadly similar. To assess this, we compute individual lee wave
generation rates $G_i$ using the method described in section 2.1. We then recompute these
generation rates using estimates of the current flow velocity from HYCOM at the same
locations as the current meters, and compare the resulting distributions.

Figure 1(a) shows a histogram of values of the ratio (meter $G_i$/model $G_i$) at each current
meter location, for a set of 351 current meters located in the bottom 10% of the ocean.
We see a distribution with a median of $10^{0.34}$ ($2.2\times$), mean of $10^{0.45}$ ($2.8\times$) and standard
deviation of $10^{0.96}$. A small number of very high ratios ($\sim \times 10^4$) have been excluded at
right on the figure but have been included in the numerical values; these will be discussed
below. A mean ratio between real and HYCOM currents of this order is consistent with
previous studies [Arbic et al., 2009; Scott et al., 2010; Wright et al., 2013]. The differences
seem initially quite large, but are comparatively small given the significant uncertainties
in the bottom currents derived from global ocean models and the vast range covered by
our estimates (results spread over 5 orders of magnitude). This assumption is justified by
the scatter plot (figure 1(b)), which shows a relatively strong spatial correlation ($r=0.67$)
between the two data sets.

We next examine the spatial distribution of these estimates using the same set of meters,
figure 2. Since we are using current meter data to generate our estimate of $G$ and using
the model data only to extrapolate the measurements out to the whole ocean via bias
correction and region determination, this spatial distribution is much more crucial to
our analysis. Figure 2(a) shows the model estimates and 2(b) our physical measurements;
coloured points indicate the locations of measurements, whilst the filled areas in 2(a) show
the model estimates at each model gridpoint and in 2(b) the current meter measurements
extrapolated outwards into regions defined by our cluster methodology, for a fixed set of
100 clusters, using the mean value for each region shown. Blue filled areas show where
we have both a physical and model-based estimate, whilst grey filled areas show where
we only have a model estimate. White regions indicate either land or regions where our
topographic roughness data suggests that the terrain is too flat to generate a significant
lee wave signal [Scott et al., 2011]. The shapes and values of extrapolated regions, such as
those shown in 2(b), differ sharply between different sets of cluster analyses, and should
be considered merely as indicative of the methodology; locations under- or over-estimated
by one set of clusters may be reversed in another.

4. Results

We next combine these regional estimates to produce a global estimate of the lee wave
generation rate $G$ as described in the Methods section; by varying the number of clusters
generated, we obtain a range of estimates of $G$. Figure 3 illustrates these results. The
main panel shows the effect of varying the number of clusters used; the grey shaded regions
indicate uncertainty bounds (light grey: 95% of distribution/2 st.dev., dark grey: 68%
of distribution/1 st.dev). The primary estimate of $G$ shows significant variability when
the number of clusters is less than approximately 100, with estimates of $G$ well above
the 1 TW of wind power input to the ocean system, and consequently clearly unphysical.
The dramatic variability at low cluster numbers is consistent with the results of Wright
et al. [2013], which were computed using a similar method. Above 100 clusters, however,
values settle down and stabilise at reasonable values. Averaging our results over analyses
with > 150 clusters and taking the widest error bounds in this range gives an estimate for
$G$ of $0.75 \pm 0.17$ TW ($\pm$ 2 st. dev.). This range is in addition to any additional unknown
uncertainty imposed by the topographic dataset, which is believed to be of similar mag-
nitude [Scott et al., 2011]. Our result is significantly higher than previous estimates, and
the consequences of this will be discussed below.

Ideally, we would consider only locations where we have a clear measurement of the
current flow velocity in the bottom layer of the ocean. However, this is tricky to define.
To assess this, the inset panel of figure 3 shows the sensitivity of this estimate to the
percentage of the ocean considered to represent bottom current flow speeds. The analysis
was performed by analysing current meters in different depth ranges from the bottom 1%
to the bottom 20%. When using only current meters in the bottom 5% or less, results are
unstable due to highly limited data coverage, but then remain stable to within uncertainty
bounds in the range 5%–12% before entering a different regime with slower currents above
this. Accordingly, using current meters in the bottom 10% provides a compromise within
this range. This sensitivity test was repeated for a range of cluster numbers, with similar
results.

We also assessed the effects of changing the $N$ climatology from one derived using the
WOA2009 dataset to one derived using the OCCA dataset, as described above. The
results of this are shown in the supplementary material for this article (figure S1). Using
the same criteria as above, we obtain an estimate of $G$ of $0.57 \pm 0.16$ TW ($\pm$ 2 st. dev.);
this estimate is lower than our primary estimate, but still larger than typical literature values.

5. Discussion and Conclusions

By comparing to HYCOM, we estimate our cluster analysis, while only covering around 65% of the geographic area of the World Ocean, covers regions contributing >95% of $G$.

At cluster numbers greater than $\sim$150, we observe the largest generation rates in regions where a strong current is combined with a rough ocean bottom; in particular, high values of $\overline{G_i}$ are observed around the Agulhas current, the Southern Atlantic, and parts of the Antarctic Circumpolar current. The contribution to the global integral from the Gulf Stream is comparatively small, despite the strong currents, due to the relatively smooth bottom topography in most of this region.

The mapped results show significant spatial homogeneity between the in situ and modelled estimates at the current meter locations (denoted by symbols); consistent with figure 1, the great majority of model estimates lie within around an order of magnitude of the measurements. The most significant divergences are in three principal locations: two sets of measurements, (1) off the western seaboard of North America and (2) off the south-eastern coast of Greenland, where model estimates are around a tenth of observed measurements and, by far the biggest divergence, a set of measurements (3) in the Atlantic Ocean southwest of the West African coast, where the model estimates are $\sim10^{-4}$ times the in situ measurements. These regions have been individually assessed; in each region, there are multiple (>5) separate locations where measurements are similar (within an order of magnitude), and the in situ time series show no evidence of any spikes or
other strong irregular divergences from the mean which could bias the result. Consequently, these results are included in our analyses; it should be noted that the African set, when extrapolated out, typically represent around 10% of the total global estimate of $G$ in our analyses. A single measurement to the south of Japan and a set of three measurements near Antarctica were similarly assessed and found to be unrepresentative of their regions and consequently have been omitted from our analysis; in particular, the single high-valued measurement off the coast of Japan if left in the analysis would have contributed around 10% of the global total $G$ in a region where model estimates are very low, dominating our global estimate.

Similar analyses were performed using the GLORYS model [Scott et al., 2012], but were found to represent the in situ data considerably more poorly than HYCOM.

Both our estimates are considerably larger than the previous model-based estimates of Scott et al. [2011] and Nikurashin et al. [2012] and also larger than, for example, the strawman estimate for wave processes of [Wunsch and Ferrari, 2004]; Scott et al. [2011] suggested that HYCOM had current speeds biased low by a factor $\sim 2$, consistent with the difference observed here. In fact, our primary estimate suggests that lee wave generation may be the dominant sink of the current energy deposited by surface winds, converting at least 50% and potentially almost all of this energy to lee waves. Since these waves can potentially propagate both considerable horizontal distances and vertically through the column to heights well above the bottom boundary layer, this process would deposit their energy for mixing at locations far away from the bottom topographic features which generate them; since altering the vertical distribution of the mixing has been found to
have an important effect on the simulated general circulation [e.g. Saenko et al., 2012],
this difference in mechanism has important potential ramifications.

Acknowledgments. CJW was funded by Stratégie d’Attractivité Durable from the
Région Bretagne awarded to RBS. RBS acknowledges funding provided by the CNRS and
by NSF grants OCE-0851457 and OCE-0960834, NASA grant number NNX10AE93G, a
contract with the National Oceanography Centre, Southampton, UK. DF was funded by
a Marie Curie Career Integration Grant awarded to RBS. Portions of the write-up of the
article was carried out during an academic visit to the University of Oxford by CJW,
generously arranged by Prof. Lesley Gray. The authors would additionally like to thank
the anonymous reviewers of this and a previous version of this article, Dr John Goff for the
topographic roughness dataset, Dr Jenny Hanafin for supplying the OCCA climatological
data, and the many groups and the many scientists who deployed the original measurement
instrumentation at various times and locations.

References

Arbic, B. K., et al. (2009), Estimates of bottom flows and bottom boundary layer dissipation
of the oceanic general circulation from global high-resolution models, Journal of

Bell, T. H. (1975), Topographically Generated Internal Waves in the Open Ocean, Journal

Bühler, O., and M. E. McIntyre (2005), Wave capture and wave-vortex duality, Journal


Scott, R. B. (1999), Geostrophic energetics and the small viscosity behaviour of an idealized ocean circulation model, Ph.d. dissertation, McGill University, Montreal, PQ,


Figure 1. (a) Histogram illustrating difference between GMACMD and HYCOM lee wave generation rate $G_i$ estimate at each current meter location. (b) Scatterplot showing the same data. The solid line indicates a least-squares regression fit to the data; the dashed line indicates a 1:1 correspondence between the two datasets.

Figure 2. Maps showing estimates of $G_i$ at (symbols) each current meter location and (contours) each point, as derived from (a) HYCOM data and (b) GMACMD current meters. Grey colour scale shows regions where no current meter data was used; blue shows regions contributing to our global estimate; values are derived from (a) the nearest HYCOM grid point and (b) the extrapolated current meter values for this set of clusters (see text for more detail).

Figure 3. Main panel: lee wave generation rate $G_{globe}$ as a function of the number of clusters, derived using the WOA2009-derived $N$ climatology. Inset panel: $G_{globe}$ as a function of the percentage of ocean depth included in the analysis, at a constant 200 clusters. Dashed lines indicate the points that correspond on each figure.

Figure 4. (Figure S1) As figure 3, but using the OCCA-derived climatology for $N$. Main panel: lee wave generation rate $G_{globe}$ as a function of the number of clusters. Inset panel: $G_{globe}$ as a function of the percentage of ocean depth included in the analysis, at a constant 200 clusters. Dashed lines indicate the points that correspond on each figure.