Fisheries and Oceans Canada / Pêches et Océans Canada

NOAA FISHERIES
NATIONAL MARINE FISHERIES SERVICE

CERT

Comité d'évaluation des ressources transfrontalières

Document de travail 2014/21

Ne pas citer sans
autorisation des auteurs

TRAC

Transboundary Resource Assessment Committee

Working Paper 2014/ 21

Not to be cited without
permission of the authors

Kriged Estimates of Yellowtail Flounder Biomass in the Closed Area II Access Area Based on the Georges Bank Pilot Flatfish Survey

Charles Adams

NOAA Fisheries, Northeast Fisheries Science Center
166 Water Street, Woods Hole, Massachusetts

Ce document est disponible sur l'Internet à :

This document is available on the Internet at :

http://www.mar.dfo-mpo.gc.ca/science/TRAC/trac.html

ABSTRACT

Kriged estimates of yellowtail flounder biomass in the Closed Area II Access Area were calculated using data from the Georges Bank pilot flatfish survey. Differences in spatial structure were identified, so estimates of biomass were derived separately for each vessel, as well as a combined estimate for both vessels. Kriged estimates of yellowtail biomass were 1,652 mt, 1,767 mt and 1,683 mt for the Mary K, Yankee Pride and combined data, respectively. These estimates were more precise than those obtained by simple swept area expansion of the arithmetic mean kg/tow.

Introduction

The objective of this working paper was to get kriged estimates of yellowtail flounder biomass in the Closed Area II Access Area (CA2AA) using standard geostatistical techniques (Rivoirard et al. 2000).

Methods/Results

Data from the Georges Bank pilot flatfish survey were used in this analysis. There were n = 32 samples for the Mary K, and n = 39 samples for the Yankee Pride within the bounds of the CA2AA (Figure 1). The two vessels were initially analyzed separately as part of exploratory data analysis. Differences in spatial structure were identified at this stage, so kriged estimates of biomass were derived separately for each vessel. A combined estimate was also calculated. In this case two nearly collocated samples caused erratic variogram behavior. Thus, these two data points were removed and the combined analysis was done on the remaining n = 69 samples.

Geographical referencing was done by setting the southwest corner of CA2AA as (0, 0) and converting all coordinates to km (Rivoirard et al. 2000). The midpoint of the latitude was used to convert longitude. Thus, kriged maps are presented in easting and northing rather than longitude and latitude.

Empirical variograms were constructed using the classical estimator of Matheron (1963):

$$\gamma(h) = \frac{1}{2N(h)} \sum [z(s) - z(s+h)]^2 \qquad (1)$$

where $\gamma(h)$ is the semivariance at distance h, $z(s)$ is the observed value of interest, in this case yellowtail weight (kg/tow), at location s. Authorized models were fit to all empirical variograms using the automated procedure of Desassis and Renard (2013) implemented in the R package RGeoS (Renard et al. 2013). Typically, model fits yield nugget, sill and range estimates. In this particular study the nugget, which is a discontinuity from the origin at distance h = 0 that describes measurement error and/or microscale variation, was not necessary to minimize the cost function. The other two parameters resulting from model fits are the sill, which is the asymptote of γ that occurs at distance $h = a$. The latter, referred to as the range, describes the extent of spatial correlation in the data (Journel and Huijbregts, 1978). Geometric anisotropy, which occurs when the range changes with direction while the sill remains constant (Journel and Huijbregts, 1978), was also modeled. This is reported here as: the major range (direction of maximum spatial continuity) at angle φ (measured clockwise from north in degrees); and the minor range, calculated at the angle perpendicular to φ.

Three biomass estimates are presented using geostatistical methods described by Rivoirard et al. (2000):

1. *Assuming no spatial correlation*. The arithmetic mean and variance are calculated with equal weighting. The coefficient of variation, assuming the sample values are independent and identically distributed, is:

$$\mathrm{CV_{iid}} = \frac{\sigma}{\bar{z}\sqrt{N}} \qquad (2)$$

2. *Arithmetic mean with geostatistical variance*. The geostatistical estimation variance is calculated as:

$$\sigma_{\mathrm{E}}^2 = 2\bar{\gamma}(z,V) - \bar{\gamma}(z,V) - \bar{\gamma}(z,z) \qquad (3)$$

   where *V* is the domain, in this case CA2AA. The coefficient of variation, using the geostatistical structure, is:

$$\mathrm{CV_{geo}} = \frac{\sigma_{\mathrm{E}}}{\bar{z}(V)} \qquad (4)$$

3. *Kriged mean and kriging variance*. A weighted mean:

$$\sum \lambda_i z_i \qquad (5)$$

   is calculated subject to the constraint that the kriging weights *λ* sum to 1. The estimation variance is the kriging variance:

$$\sigma_{E}^2 = 2\sum \lambda_i \bar{\gamma}(z,V) - \bar{\gamma}(V,V) - \sum_i \sum_j \lambda_i \lambda_j \gamma(z_i - z_j) \qquad (5)$$

   and the coefficient of variation is calculated according to Eq. 4.

Estimates derived with methods #2 and #3 used a mesh size of 100 × 100 to discretize the domain *V*, as further increases in mesh size resulted in only trivial improvements (i.e., four decimal places) in $\mathrm{CV_{geo}}$ (Journel and Huijbregts 1978; Rivoirard et al. 2000). Kriged maps are presented for estimates derived from method #3.

Data for the Mary K were best fit by a cubic model with sill = 866.9, major range = 24.0 km, minor range = 11.0 km, and anisotropy angle of 110.0° (Figure 2). Data for the Yankee Pride were best fit with nested spherical (sill = 688.2, major range = 21.0, minor range = 19.4) and linear (slope = 2.1) models containing an anisotropy angle of 131.5° Figure 3). The combined data were best fit with nested cubic (sill = 563.6, major range = 22.4 km, minor range = 4.0 km) and linear models (slope = 10.6) containing an anisotropy angle of 100.8° (Figure 4). Note that a linear model does not indicate a drift (large-scale trend) unless λ increases faster than h^2 (Journel and Huijbregts, 1978; Rivoirard et al., 2000; Webster and Oliver, 2007).

Biomass estimates based on the arithmetic mean were 1,519 mt, 1,794 mt and 1,712 mt for the Mary K, Yankee Pride and combined data, respectively (Table 1). Comparison of CV_{iid} vs. CV_{geo} shows only minor improvements in precision by incorporating the spatial structure into the variance for the Mary K and Yankee Pride, respectively. However, CV_{geo} for the combined data was higher than CV_{iid}. Given the quasi systematic study design, this indicates that the CV_{iid} for all three estimates are incorrect.

CV_{geo} was substantially lower for the kriged means, indicating that these biomass estimates were more precise. In the case of the Mary K, a 0.035 improvement in CV_{geo} was accompanied by a 132 mt increase in the biomass estimate; in the case of the Yankee Pride a 0.026 improvement in CV_{geo} corresponded to a 27 mt decrease in the biomass estimate; and for the combined data a 0.029 improvement in CV_{geo} was accompanied by a 30 mt decrease in the biomass estimate.

Kriged maps for the Mary K (Figure 5), Yankee Pride (Figure 6) and the combined data (Figure 7) visualize spatial differences in predicted yellowtail (kg/tow) between the three estimates.

Discussion/Conclusions/Summary/Recommendations

Samples separated by distances less than the range are autocorrelated (Journel and Huijbregts 1978). The major range for the Mary K, Yankee Pride and combined data were between 21 – 24 km, indicating that samples separated by distances less than this along the NW – SE gradient are not independent. However, the scale of the structure was notably different: for the Mary K it could be described by a standard anisotropy ellipse with major and minor axes of 24 km and 11 km, respectively; whereas for the Yankee Pride it was an almost isotropic structure seemingly nested within a larger scale structure stretching into the so-called “donut hole” of Georges Bank; this is clearly an artefact of the two large hauls near the northern boundary of the CA2AA.

Kriged yellowtail biomass estimates for CA2AA in this analysis were 1,652 mt, 1,767 mt and 1,683 mt for the Mary K, Yankee Pride and combined data, respectively. These values are near the higher end of the range of estimates in working paper #22, which were also calculated for the CA2AA. Biomass estimates described in this paper illustrate the improved precision associated with kriging relative to the standard swept area expansion of the arithmetic mean kg/tow, which does not take into account the underlying spatial structure of the data.

Literature Cited

Desassis N, Renard D. 2013. Automatic variogram modelling by iterative least squares: univariate and multivariate cases. Mathematical Geosciences 45:453–470.

Journel AG, Huijbregts CJ. 1978. Mining Geostatistics. Academic Press, London.

Matheron G. 1963. Principles of geostatistics. Economic Geology 58:1246–1266.

Renard D, Bez N, Desassis N, Beucher H, Ors F, Laporte F. 2013. RGeoS v9.1.10. Mines ParisTech: www.cg.ensmp.fr/rgeos.

Rivoirard J, Simmonds J, Foote KG, Fernandes P, Bez N. 2000. Geostatistics for Estimating Fish Abundance. Blackwell Science, Oxford.

Tables

Table 1. Summary statistics and geostatistical estimation results for yellowtail flounder by vessel: area of CA2AA (km^2), arithmetic mean kg/tow, standard deviation σ, coefficient of variation CV and associated biomass estimate; coefficient of variation of the mean assuming no spatial structure CV_{iid}, estimation variance σ_E, and the coefficient of variation of the uncertainty for the estimation of biomass within CA2AA using the geostatistical structure CV_{geo}; kriged mean kg/tow, estimation variance σ_E, coefficient of variation of the uncertainty for the kriged estimate of biomass within CA2AA using the geostatistical structure CV_{geo} and associated biomass estimate.

		Unweighted							Kriged			
Vessel	Area (km^2)	Mean (kg/tow)	Biomass (mt)	σ	CV_{sam}	CV_{iid}	σ_E	CV_{geo}	Mean (kg/tow)	Biomass (mt)	σ_E	CV_{geo}
Mary K	3,880	28.7	1,519	29.9	1.04	0.184	5.2	0.181	31.3	1,652	4.5	0.145
Yankee Pride	3,880	33.9	1,794	32.0	0.94	0.151	4.9	0.144	33.4	1,767	3.9	0.118
Combined	3,880	32.4	1,712	31.3	0.97	0.116	4.2	0.129	31.8	1,683	3.2	0.101

Figures

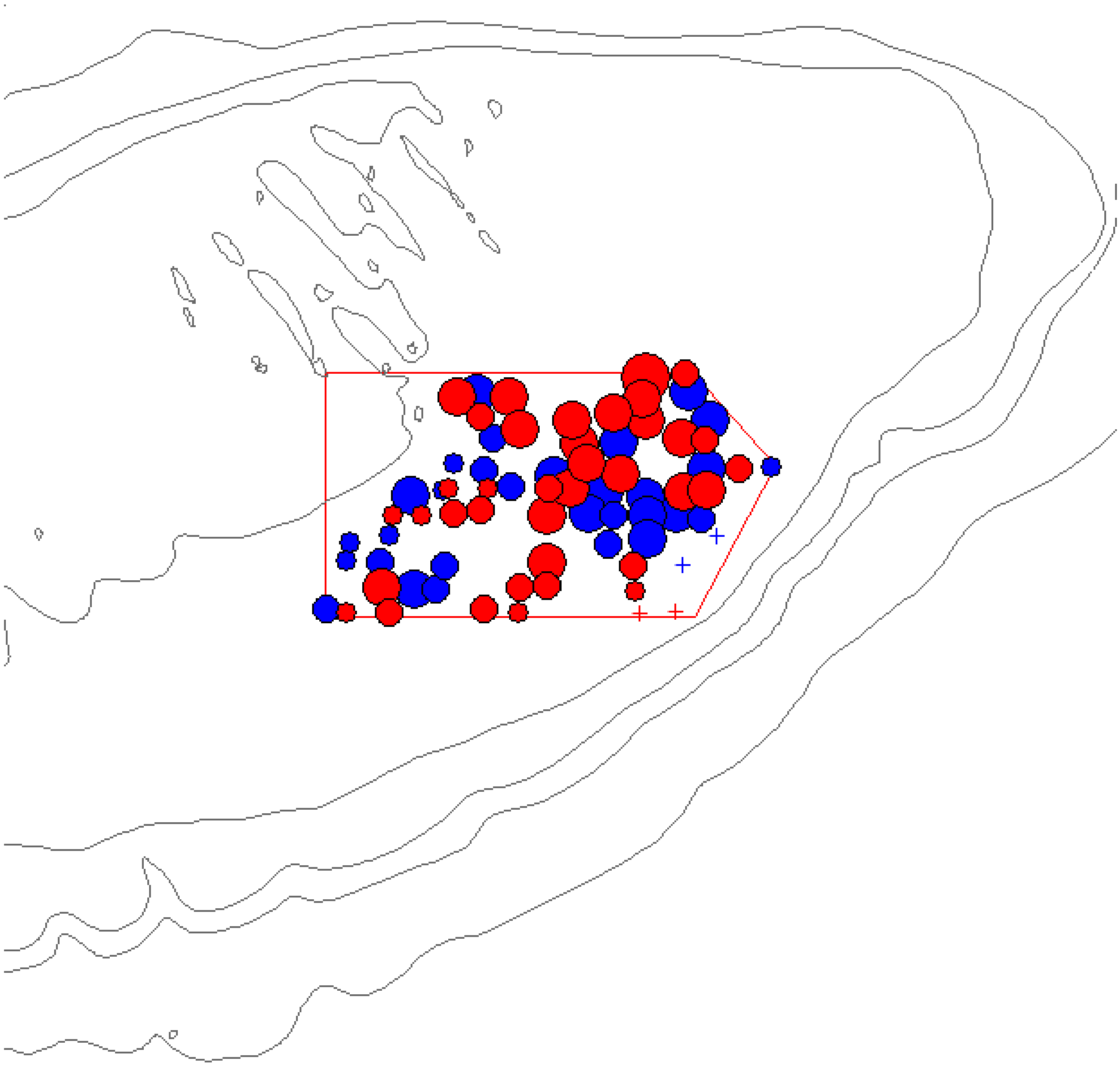

Figure 1. Map showing the boundaries of CA2AA on Georges Bank, and geographic midpoints of tow locations for the Mary K (blue) and Yankee Pride (red) used in this analysis. Size of circle is proportional to yellowtail weight (kg). Crosses indicate stations where no yellowtail were caught.

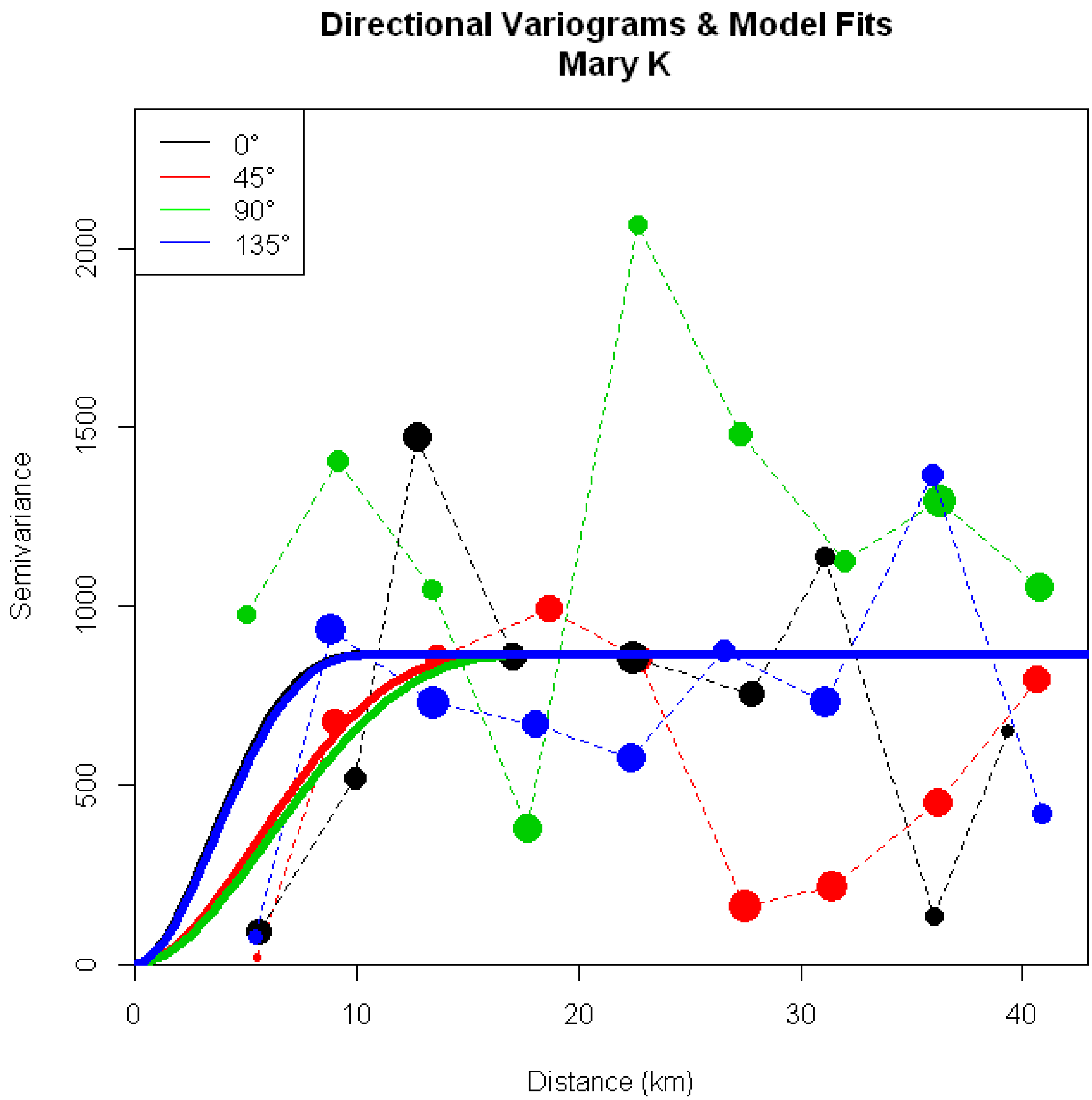

Figure 2. Empirical variograms and theoretical model fits of yellowtail flounder weight (kg/tow) for the Mary K. Directional variograms are shown at the standard exploratory angles for comparative purposes.

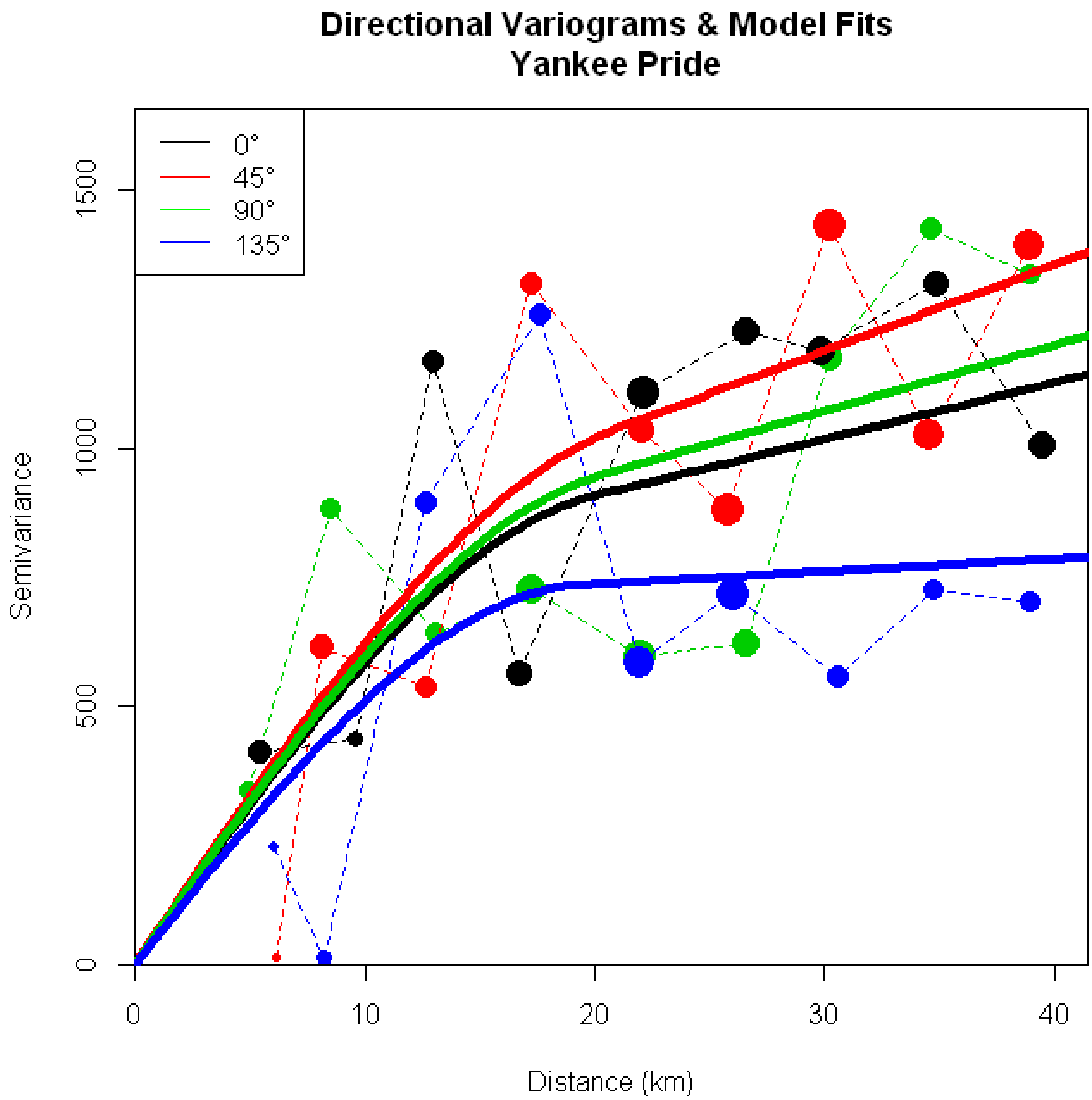

Figure 3. Empirical variograms and theoretical model fits of yellowtail flounder weight (kg/tow) for the Yankee Pride. Directional variograms are shown at the standard exploratory angles for comparative purposes.

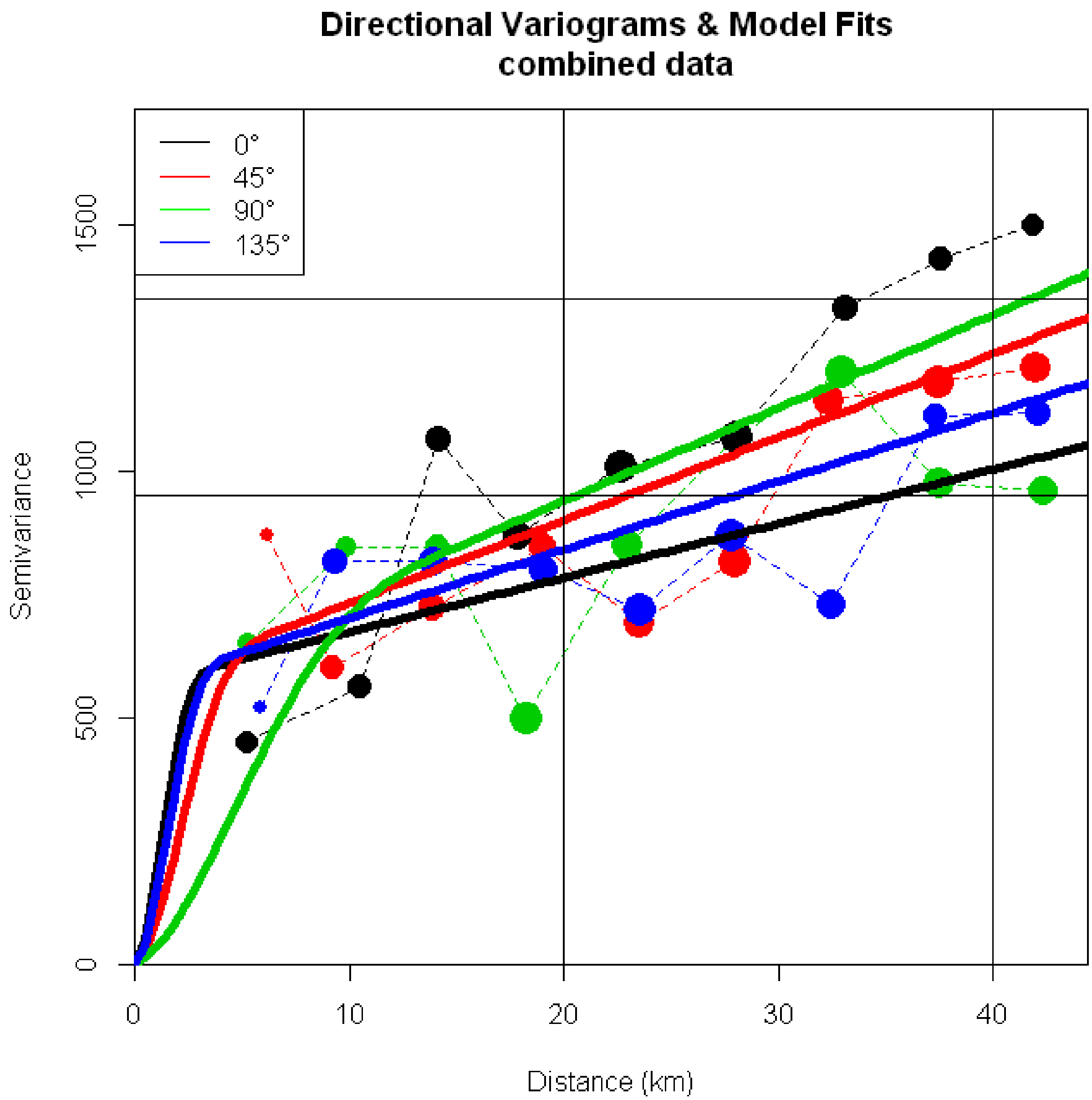

Figure 4. Empirical variograms and theoretical model fits of yellowtail flounder weight (kg/tow) for the combined data. Directional variograms are shown at the standard exploratory angles for comparative purposes. Reference lines at semivariance = 950 and 1350 illustrate that λ is not increasing faster than km^2.

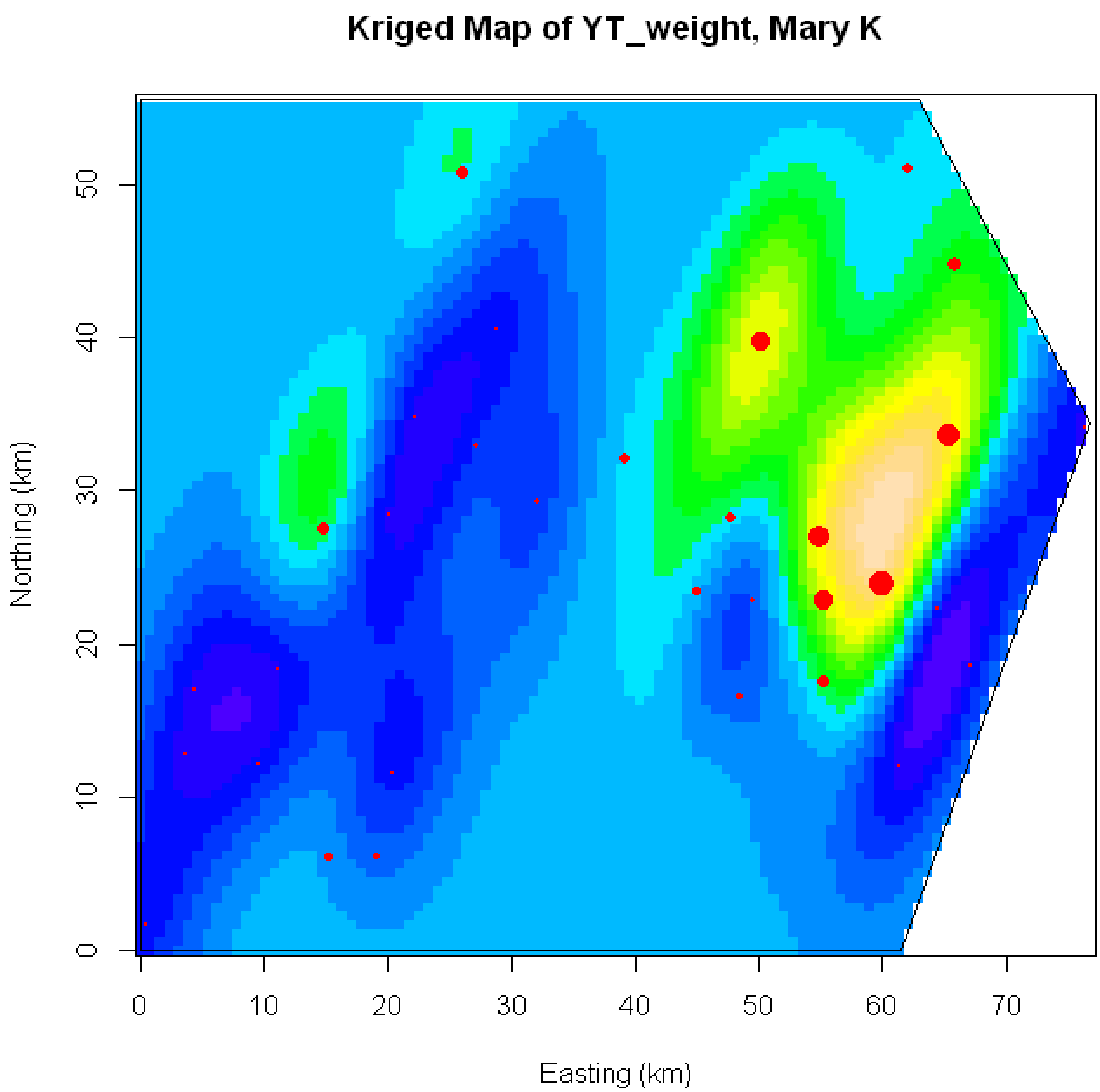

Figure 5. Kriged map of yellowtail flounder weight (kg/tow) for the Mary K. Size of red circle is proportional to yellowtail weight (kg/tow).

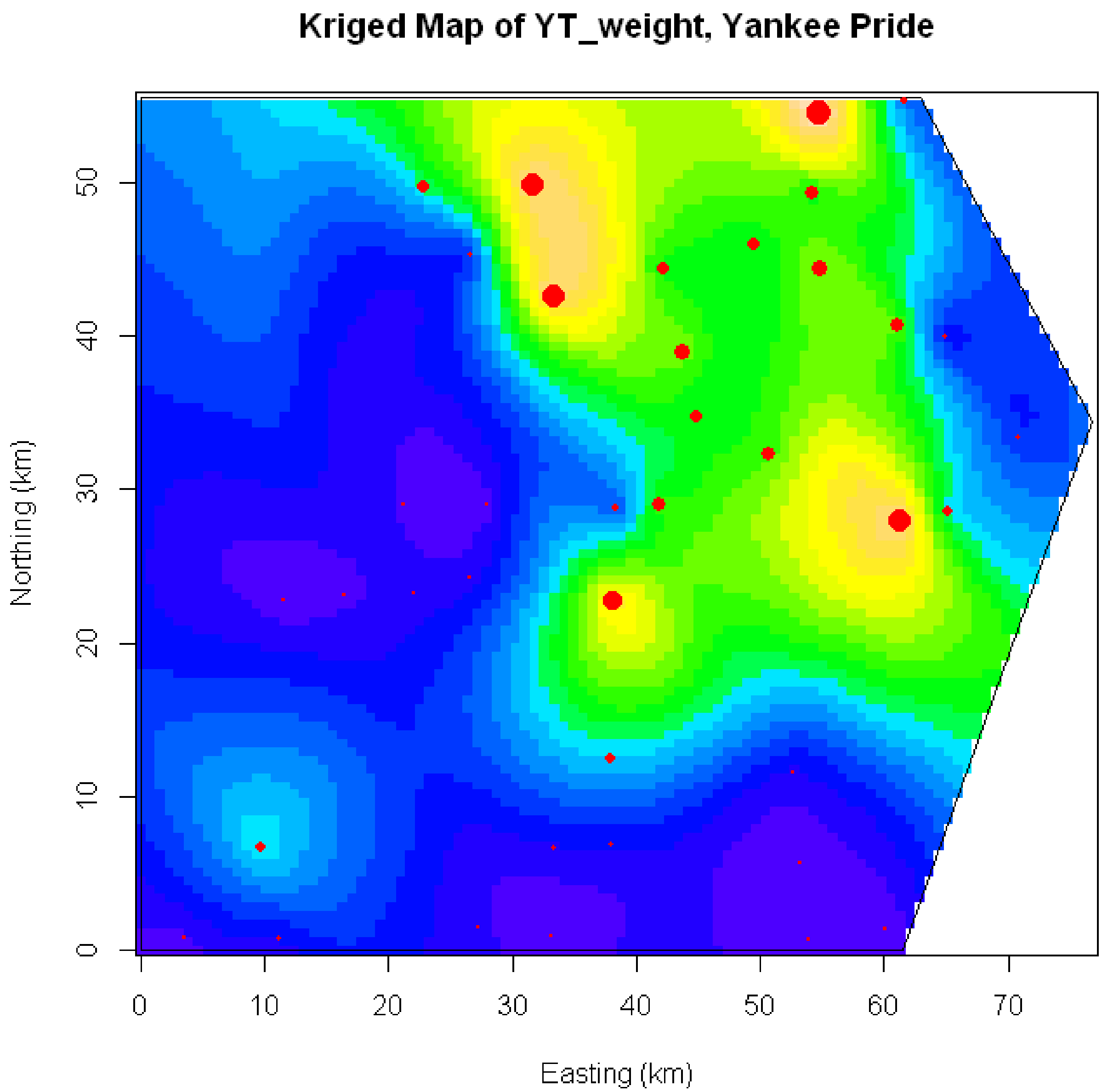

Figure 6. Kriged map of yellowtail flounder weight (kg/tow) for the Mary K. Size of red circle is proportional to yellowtail weight (kg/tow).

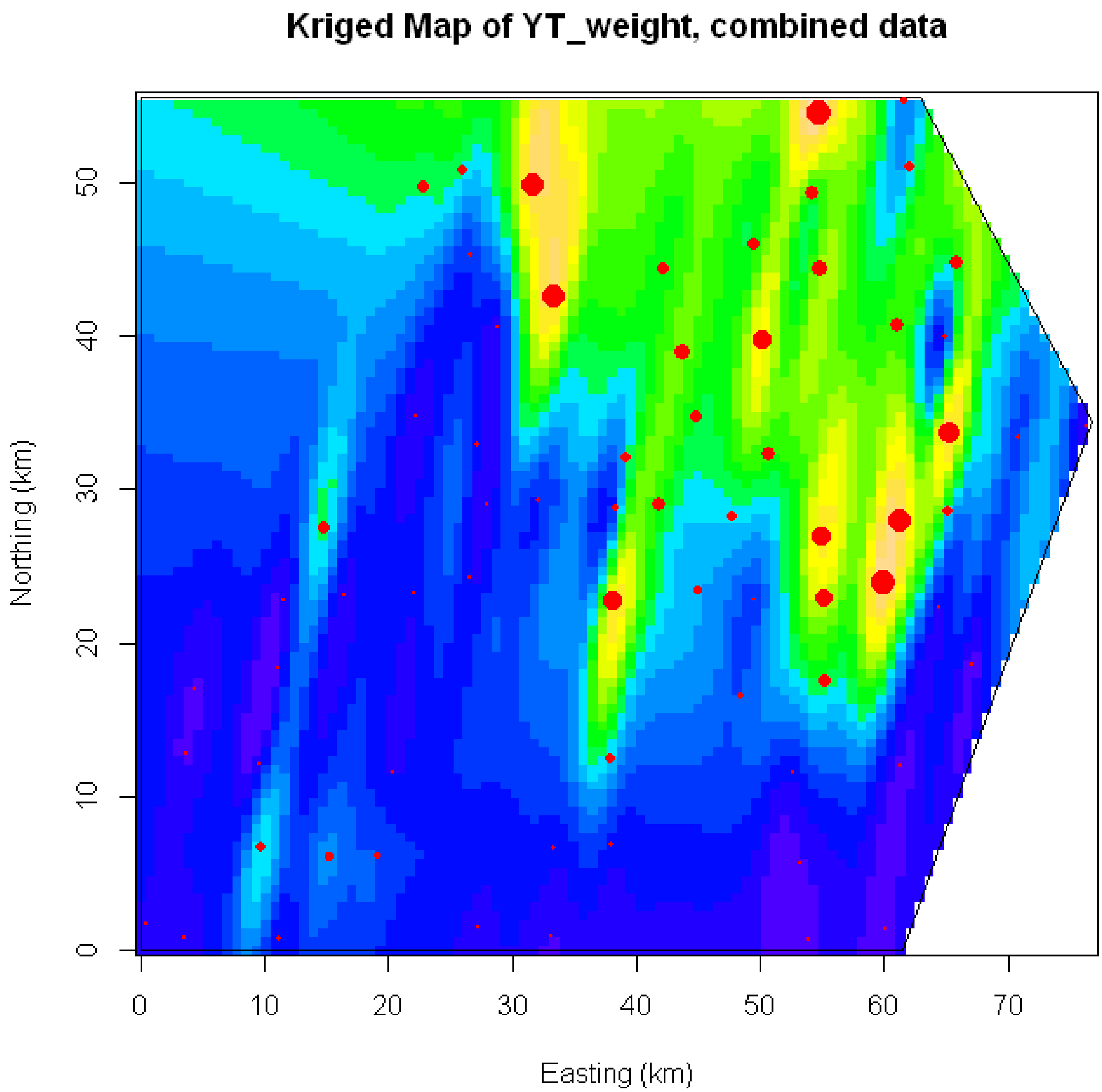

Figure 7. Kriged map of yellowtail flounder weight (kg/tow) for the combined data. Size of red circle is proportional to yellowtail weight (kg/tow).