LHCb/2007-011
OFFLINE
September 19, 2007

Primary vertex reconstruction

M. Krasowski, M. Kucharczyk, W. Manner, G. Polok, M. Witek

Institute of Nuclear Physics PAN, Krakow, Poland

Abstract

This report reviews the improved method of primary vertex off-line reconstruction. The new method for primary vertex seed finding as well as modified procedure for vertex fit have been briefly described. Results of the primary vertex reconstruction performance are presented and compared to the previous version of the package. Improvement in the primary vertex reconstruction efficiencies is observed. Small but statistically significant bias, coming mostly from the long-living tracks as well as from a bias of the Velo measurements present in the DC06 simulation, affects the determination of the primary vertex position along the beam direction.

Contents

1 Introduction **2**

2 Primary vertex seeding **2**

3 Fitting procedure **3**

4 Performance of primary vertex reconstruction **3**

4.1 Reconstruction efficiency and the rate of false PV's 3

4.2 Primary vertex resolution 6

4.3 Study of bias on primary vertex z position 6

5 Conclusions **9**

1 Introduction

There are three kinds of primary vertex (PV) finding procedures, two of them at the trigger level [1, 2] and one for the off-line analysis. In the present note we concentrate on the off-line primary vertex reconstruction.

There are two main steps of the primary vertex reconstruction. The first one is to search for PV seeds providing the set of z coordinates of PV candidates. In the second step PV reconstruction is performed relying on the iterative fit using tracks extrapolated to the z-coordinate of PV seed.

The primary vertex algorithms have beed adapted for the new MC data release (DC06) and new track event model. New matrix algebra from the MathCore package has been applied instead of CLHEP package.

The results shown in the present note have been prepared using a sample of about 5.0×10^4 inclusive $b\bar{b}$ events reconstructed by Brunel v30r14 for DC06. The comparison to DC04 performance was based on the data reconstructed by Brunel v29r11.

2 Primary vertex seeding

The initial estimate of the primary vertex z coordinate is based on the seeding procedure employing the method of analytical clusterization[1]. It turned out that more emphasis should be put on the tracks with large angles θ with respect to the z axis as the error σ_z along the beam is proportional to $\sigma_{d_0}/sin\theta$, where d_0 stands for impact parameter and σ_{d_0} denotes its error. The cluster is defined by the z^{clu} coordinate and its uncertainty σ_z^{clu}. The procedure starts with initial clusters determined by the closest approaches of the tracks with respect to the z axis as well as the errors assigned to them. The general principle of the method is as follows. In each iteration a list of all pairs of clusters are considered, and a single pair with minimum distance is selected. The distance itself is defined as:

$$D^{pair} = |z^{clu_1} - z^{clu_2}|/\sqrt{(\sigma_z^{clu_1})^2 + (\sigma_z^{clu_2})^2}. \quad (1)$$

Selected pair is merged into one cluster under the condition that $D^{pair} < 5$. The z^{clu} as well as σ_z^{clu} of the merged cluster are taken as a weighted mean of the z coordinates of two initial clusters, i.e.

$$z^{clu} = \frac{1}{w} \cdot (w_1 \cdot z^{clu_1} + w_2 \cdot z^{clu_2}), \quad (2)$$

$$\sigma_z^{clu} = 1/\sqrt{w}, \quad (3)$$

where $w_i = (1/\sigma_z^{clu_i})^2$ $(i = 1, 2)$ and $w = w_1 + w_2$. Then the next iteration is performed. The procedure stops if no more pair of clusters remains to be merged. The final clusters are the candidates for primary vertex seeds. In the final step the

[1]The previous version of the primary vertex reconstruction package relied on the histogramming and peak finding method.

quality of such candidates is checked. Clusters with the multiplicity (number of tracks the cluster consists of) below 6 are removed. Additional quality conditions are also applied to reduce the rate of false clusters. Main source of false seeds comes from clusters with relatively low multiplicities, which are formed around tracks with small errors pointing close to the high-multiplicity primary vertex. Stable working point has been found, balancing between high efficiency and the rate of false seeds.

3 Fitting procedure

For each PV seed the fitting procedure is performed in decreasing order of the seed multiplicities. The Long, Upstream as well as Velo-only tracks are used. They are extrapolated to the position of the seed. Long and Upsteram tracks are extrapolated using full transport service while Velo-only tracks are extrapolated as straight lines. In the present note only main idea of the fitting procedure is described. More details can be found in [3]. The position of the primary vertex is determined by the least square method minimizing following χ^2:

$$\chi^2_{PV} = \sum_{i=1}^{n_{tracks}} \frac{d^2_{0_i}}{\sigma^2_{d_{0_i}}}, \tag{4}$$

where d_0 denotes the impact parameter of a track while σ_{d_0} its error. Estimation of the PV position is determined iteratively. Initial selection of the tracks is based on a z coordinate of their closest approach to the z axis. Relatively wide window ± 30 mm around z of the seed is used.

In each iteration the new position of PV is determined. The tracks are extrapolated to the z coordinate of new PV position and the impact parameter significance, IPS $= d_0/\sigma_{d_0}$, is estimated. The track with the highest IPS is removed from the set of tracks with IPS > 4 and then next iteration is performed. The procedure stops if there is no more track to discard. In the final phase the quality conditions are imposed. PV candidates with multiplicity (number of reconstructed tracks coming from the vertex) below 6 are removed.

The whole procedure is started for the next seed excluding tracks already used by any previous PV.

4 Performance of primary vertex reconstruction

4.1 Reconstruction efficiency and the rate of false PV's

Primary vertex reconstruction efficiency is defined as a ratio of the number of reconstructed PV's and the number of reconstructible PV's. Reconstructible primary vertex is the vertex which contains at least 6 reconstructed Velo tracks corresponding to the Monte Carlo particle coming from a given primary vertex. It is useful to divide PV's into two subclasses: isolated reconstructible PV's and close reconstructible PV's as well as to measure the rate of false reconstructed PV's. The corresponding definitions are the following:

- Reconstructible PV is called isolated if its distance to the closest reconstructible PV $|\Delta z| > 10$ mm.
- Reconstructible PV is called close if its distance to the closest reconstructible PV $|\Delta z| < 10$ mm.
- Reconstructed PV is defined as false if its distance to any reconstructible PV $|\Delta z| > 5\sigma_z$, where σ_z denotes the estimated position error along z axis. Large fraction of such vertices represent reconstructed PV's corresponding to real Monte Carlo PV's, which contains reconstructed tracks, but do not fulfil the conditions to be reconstructible.

Reconstruction efficiencies as a function of PV multiplicity are shown in Fig.2. Better efficiency for the lower PV multiplicities is clearly seen for DC06. Since the shapes of track multiplicity distributions for new (DC06) and old (DC04) Monte Carlo samples used in this study are not significantly different (see Fig.1), one can compare the overall efficiencies. The results are listed in Table 1.

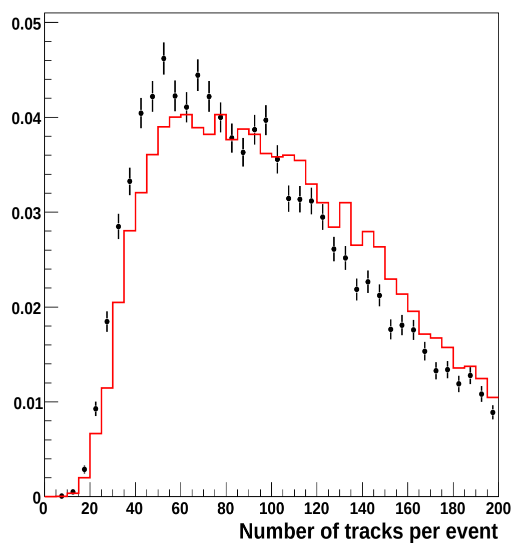

Figure 1: Comparison of track multipilcity distributions for DC06 (points) and DC04 (solid line) samples.

Lower values for DC04 sample presented in Table 1 come mostly from low multiplicity PV's. This difference is mainly caused by two effects: i) more efficient seeding procedure for DC06 version and ii) additional requirement in the reconstruction algorithm for DC04 causing the stop of primary vertex search if less than 10% of all tracks remains not assigned to any PV vertex.

The primary vertex reconstruction efficiencies for vertices with highest and second highest multiplicity are presented in Tables 2 and 3 respectively.

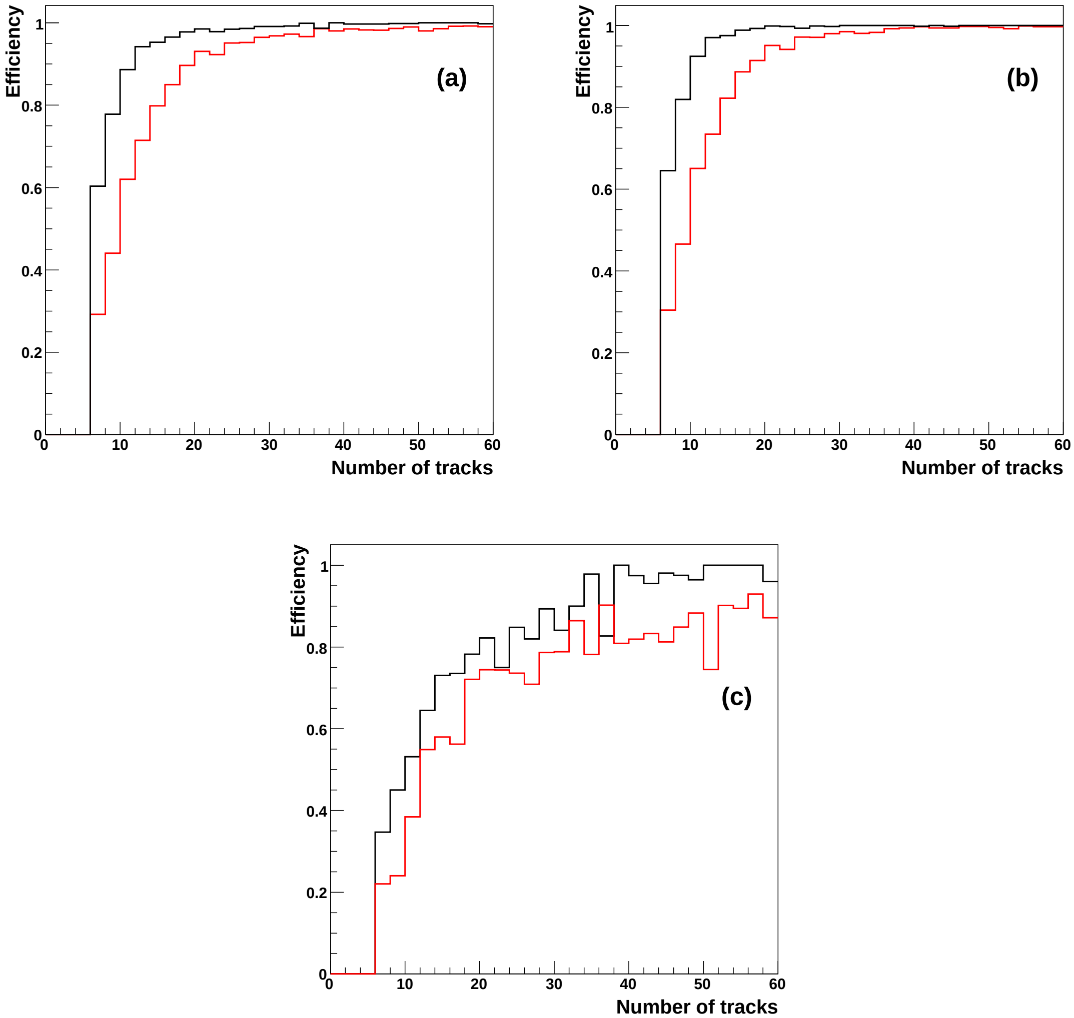

Figure 2: Primary vertex reconstruction efficiencies as a function of primary vertex multiplicity for: a) all reconstructible; b) isolated and c) not isolated PV's obtained for DC06 (black line) and DC04 (red or ligther line) samples.

Table 1: Reconstruction efficiencies for different categories of reconstructible PV's obtained for DC06 and DC04 samples.

Efficiency [%]	DC04	DC06
All reconstructible	90.0	97.5
Isolated	92.0	98.4
Close	69.7	85.6
False	1.2	0.9

Table 2: Reconstruction efficiencies for different categories of reconstructible PV's with highest multiplicity obtained for DC06 and DC04 samples.

Efficiency [%]	DC04	DC06
All reconstructible	99.5	99.6
Isolated	99.6	99.6
Close	96.2	98.8

Table 3: Reconstruction efficiencies for different categories of reconstructible PV's with second highest multiplicity obtained for DC06 and DC04 samples.

Efficiency [%]	DC04	DC06
All reconstructible	76.6	93.3
Isolated	79.0	95.6
Close	59.9	78.5

4.2 Primary vertex resolution

Spatial resolution of the primary vertex has been determined by looking at the difference of positional coordinates between Monte Carlo primary vertex and reconstructed primary vertex.

Resolutions for DC06 sample are shown in Fig.3. The values of σ_x, σ_y and σ_z are compared in Table 4 for DC06 and DC04[2]. One can see that the mean value of the distribution for Δz differs significantly from zero. Origin of this shift is discussed in the next section.

Table 4: Primary vertex resolutions σ_x, σ_y and σ_z for DC06 and DC04 samples. The values denote fit results employing single Gaussian distribution.

Resolution [μm]	DC04	DC06
σ_x	8	8
σ_y	10	10
σ_z	50	59

One can notice that results for old and new data samples do not differ much, slightly worse resolutions may be observed for DC06.

4.3 Study of bias on primary vertex z position

The mean value as well as standard deviation of Δz as a function of primary vertex multiplicity for DC06 and DC04 samples are presented in Fig.4. A statistically

[2] As has been mentioned before track multiplicity distributions for new (DC06) and old (DC04) samples are similar, thus the comparison between overall resolutions for DC06 and DC04 is justified.

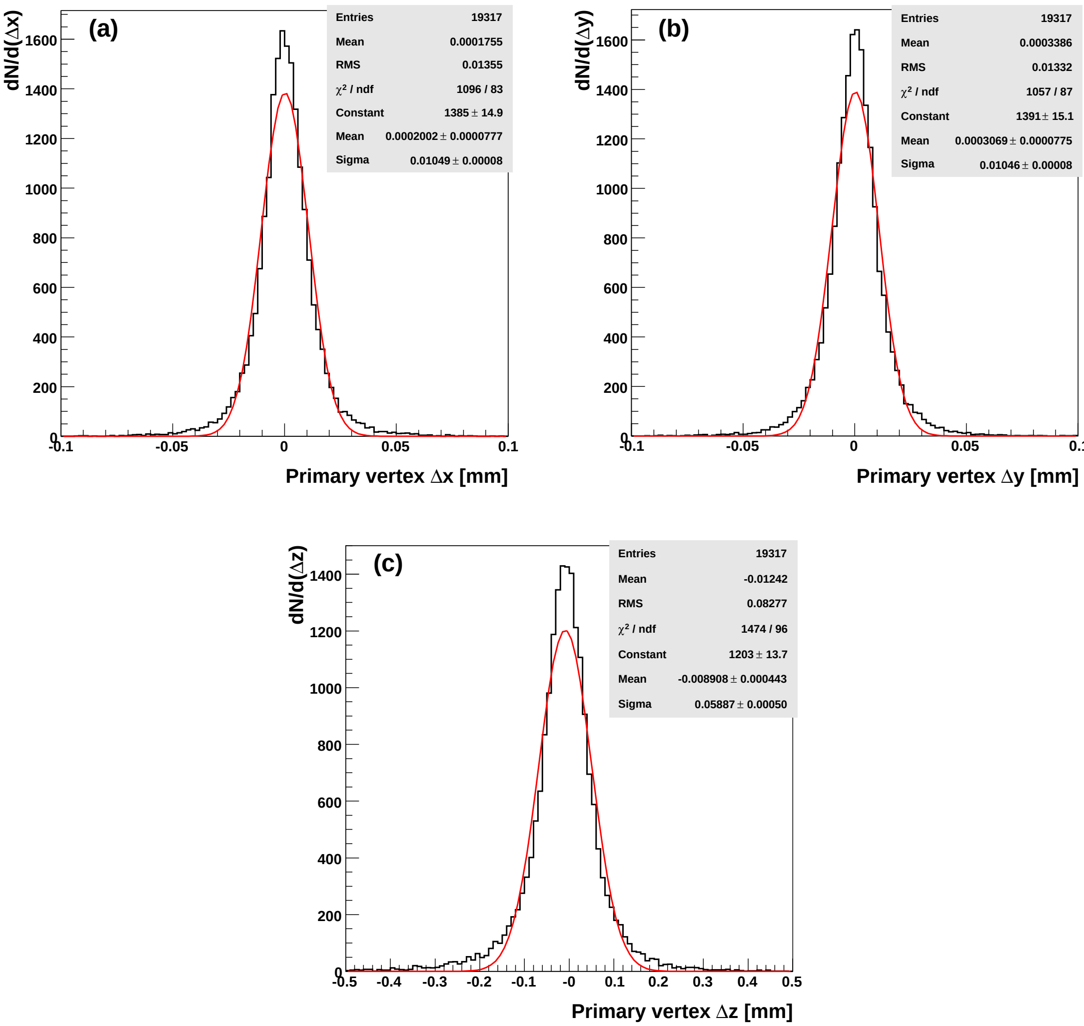

Figure 3: Primary vertex resolutions for DC06 sample. A fit to a single Gaussian is used.

significant bias in the mean value of Δz is clearly visible for both DC06 and DC04 (Fig.4a). Its meaningful contribution results from the admixture of long-living tracks that come close to the primary vertex and therefore cannot be removed by the cut on their impact parameter significance. The bias for DC06 consists of a second component caused by the Velo RZ cluster position bias.

In Fig.4b it may be observed that in overall primary vertex multiplicity range the resolution for DC06 is worse than for DC04.

In order to exclude any potential distortion of the vertex reconstruction alone, the correctness of the fitting procedure was examined in two independent ways. In order to check the algebra of the fit the tracks used for primary vertex reconstruction were shifted artificialy to an ideal position of MC truth vertex. The result of this cross-check is shown in Fig. 5, where no significant bias in Δz distribution may be observed. Moreover, stability of this procedure was confirmed by varying seed z

coordinates around its nominal value.

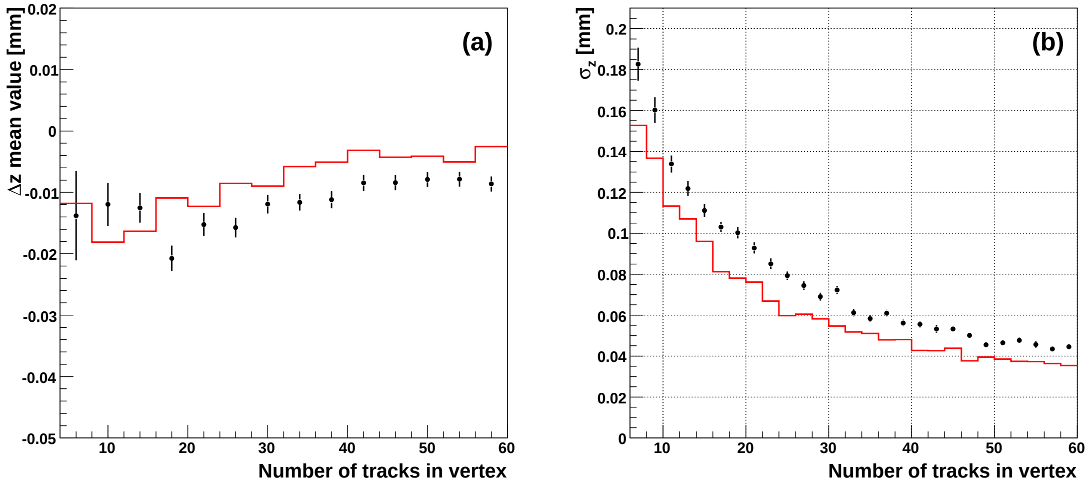

Figure 4: (a) Comparison between mean values of Δz as a function of primary vertex multiplicity for DC06 (points) and DC04 (solid line) samples; (b) Resolution of primary vertex z coordinate as a function of primary vertex multiplicity for DC06 (points) and DC04 (solid line) samples.

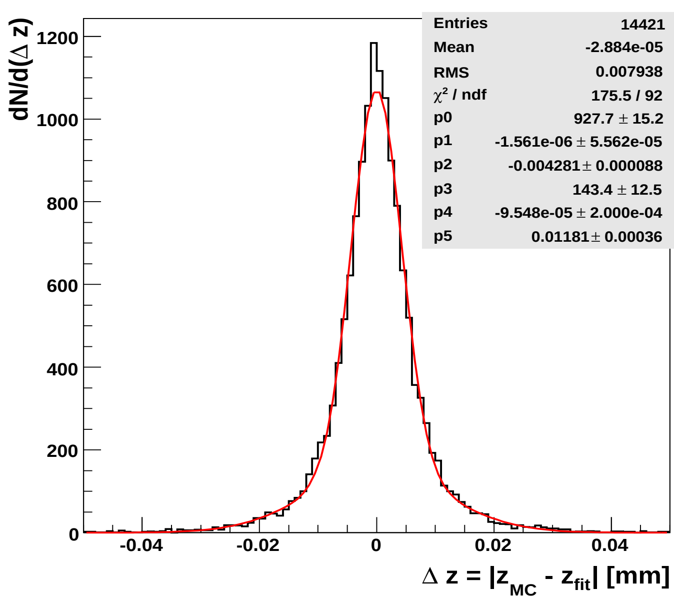

Figure 5: The result of the fit (solid line) to Δz distribution, where the tracks used for primary vertex reconstruction were coming directly from an ideal position of MC truth vertex.

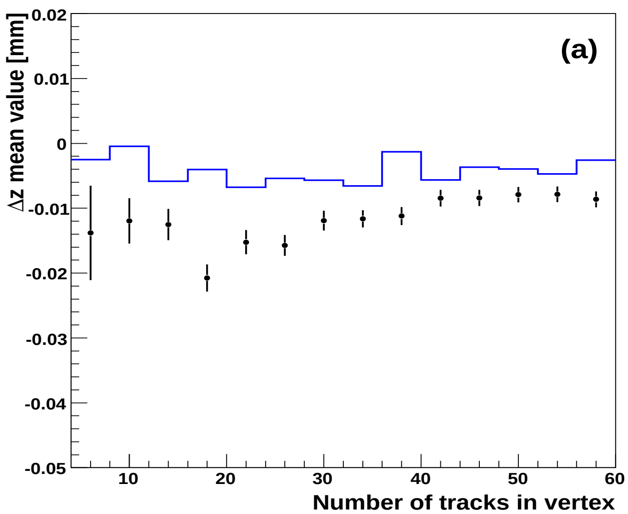

Figure 6: Comparison between mean value of Δz as a function of primary vertex multiplicity for DC06 with (points) and without (solid line) long-living tracks.

The next check was to remove the bias coming from long-living tracks by requiring that primary vertex is fitted using only reconstructed tracks positively matched to the Monte Carlo particles coming directly from the MC primary vertex. The result of this exercise is shown in Fig.6. A statistically significant shift of about 3-4 μ is clearly visible for DC06, resulting from the Velo RZ cluster position bias, while in the case of DC04 such effect is not present.

5 Conclusions

The efficiencies as well as spatial resolution of primary vertex reconstruction were compared for two different Monte Carlo data samples: DC06 and DC04. For DC06 the primary vertex algorithms have beed adapted to new track event model as well as to new matrix algebra. New effective seeding procedure based on analytical clusterization helped considerably with improving primary vertex reconstruction efficiencies, in particular for low multiplicity PV's. Rate of the false reconstructed PV's is below 1%. Slightly worse z coordinate resolution may be observed for DC06.

About 10 μm bias may be observed for both DC04 and DC06. Its main contribution comes from long-living particles. This cannot be removed by employing a simple cut on track impact parameter significance. For DC06 an addition source of the shift due to bias of the Velo measurements present in the DC06 simulation was determined to be about 4 μm.

PV reconstruction procedure described above is planned to be further developed. In particular, the functionality of refitting PV's excluding long-living particles will be examined. Furthermore, the PV reconstruction software as well as Velo simulation package are being improved. Performance of the PV reconstruction will be then updated within the subsequent note based on appropriate new data production.

References

[1] *The LHCb Level-1 Trigger: Architecture, Prototype, Simulation and Algorithm*, V. Lindenstruth *et al.*, Note LHCb-2003-064, Chapter 5.

[2] *LHCb Trigger System Technical Design Report*, The LHCb Collaboration, R.A Nobrega *et al.*, CERN/LHCC 2003-031, LHCb TDR 10, 2003.

[3] *The LHCb Level-2 Trigger*, F. Teubert, I.R. Tomalin, J. Holt , Note LHCb-98-47.