Short-term load forecasting at the distribution level predicts the load of substations, feeders, transformers, and possibly customers from half an hour to one week ahead. Effective forecasting is important for the planning and operation of distribution systems. The problem, however, is difficult in view of complicated load features, the large number of distribution-level nodes, and possible switching operations. In this paper, a new forecasting approach within the hierarchical structure is presented to solve these difficulties. Load of the root node at any user-defined subtree is first forecast by a wavelet neural network with appropriate inputs. Child nodes categorized as “regular” and “irregular” based on load pattern similarities are then forecast separately. Load of a regular child node is simply forecast as the proportion from the parent node load forecast while the load of an irregular child node is forecast by an individual neural network model. Switching operation detection and follow-up adjustments are also performed to capture abnormal changes and improve the forecasting accuracy. This new approach captures load characteristics of nodes at different levels, takes advantage of pattern similarities between a parent node and its child nodes, detects abnormalities, and provides high quality forecasts as demonstrated by two practical datasets.