

# K-means および PSO モデルアルゴリズムに基づく最適中国物流センター拠点選定の基本研究

日大生産工(院) ○劉 智琦  
日大生産工 豊谷 純

## 1. はじめに

過去 10 年間、中国の物流業界総額は 2012 年の約 177.3 兆元から 2023 年の約 352.4 兆元へと成長し<sup>1)</sup>、物流業界が国民経済の支柱として重要な役割を果たしていることを示している。2024 年の全国社会物流総額は 360.6 兆元で、価格換算では前年同期比 5.8% 増となり、増加率は前年より 0.6 ポイント上昇した。

物流需要の多様化と複雑化が進む中、全国的な物流センターの科学的な立地選定の必要性がますます高まっている。一方で、物流センターは物流ネットワークの重要な節点として機能し、その合理的な配置は物流コストの削減、輸送ルート最適化、サービス効率の向上に大きく寄与する。さらに、全国規模での物流センターの選定は、地域経済の差異に適応し、資源配分を調整し、地域間の物流の均衡ある発展を実現するためにも不可欠である。全国的な物流センターの立地選定は、産業や市場の需要を満たすだけでなく、電子商取引の急速な発展やグローバルサプライチェーンの高度化といった新たな課題にも対応する必要がある。

## 2. 先行研究

物流拠点や配送センターの選定に関する研究は、これまで多様な観点から行われてきた。

「Accessibility-Based Location of International Multimodal Logistics Hubs: A case in China」<sup>2)</sup>では、鉄道・道路・港湾・航空などの複合輸送手段の接続性を定量化することで、中国国内の国際物流ハブの到達性を評価し、重慶・鄭州・上海といった主要都市を中核拠点として特定している。この研究は交通ネットワークの到達性に着目しており、国際物流回廊の戦略的重要性を明らかにしたが、経済指標や需要側の要素は十分に考慮されていない。

一方、SALANGAM S. 「最適グローバルハブネットワークシステムの選定に関する基本研究」<sup>3)</sup>では、グローバル規模でのハブネットワークシステムの最適化問題を扱い、国際的なサプライチェーンにおけるハブ配置とネットワーク構造の効率性を探求している。しかし、そ

の研究範囲は国際レベルにとどまり、単一国家内部における物流ネットワークの実証分析には踏み込んでいない。

さらに、方法論的には、既存研究において K-means と粒子群最適化 (PSO) を組み合わせたクラスタリング手法が提案されており、クラスタ中心の初期値選択や局所解回避に有効であることが報告されている「Canonical PSO Based K-Means Clustering Approach for Real Datasets」<sup>4)</sup>。これらは画像処理やテキストクラスタリング、顧客セグメンテーション等で応用されているが、物流配送センターの選定問題に対して適用された事例は見られず、本研究の新規性を主張できる。

以上の既存研究を踏まえ、本研究の革新点は以下の通りである。研究対象の革新：国際回廊やグローバルなハブネットワークではなく、中国全土 31 省市を対象とした国家レベルの物流配送センター選定を行い、より現実的な国内物流ネットワークの最適化を目指す。指標体系の革新：交通・地理的要因に加え、GDP・人口・貨物流量といったマクロ経済指標を導入し、経済属性と空間属性を融合させた包括的なモデルを構築する。方法論の革新：従来の K-means の限界を克服するために、K-means と PSO を組み合わせたハイブリッド手法を提案し、さらに重心法を比較対象として導入することで、多手法による総合的な評価枠組みを提示する。

## 3. 数理モデル

本研究は、中国 31 省・直轄市を対象に物流配送センターの最適立地を検討し、K-means クラスタリングと粒子群最適化 (PSO) を融合した手法を提案する。人口・GDP・貨物輸送量・緯度・経度を指標として標準化したデータを用い、K-means (粗粒度  $k=6$ 、精細粒度  $k=16$ ) で初期クラスタを生成し、その質心を PSO (50 粒子・100 ステップ) で全局探索することで、局所最適を回避し総輸送コストを 40% 以上削減した。さらに、拠点間に道路距離・時間・通行料に経済罰則項を加えた複合エッジ重みを設定し、NetworkX で最小生成木 (MST) を算

出した。その結果、東中部の高需要幹線が優先され、西北・高原地域は末端枝として整理され、投資重点が明確化された。提案手法は、①K-means による分群、②PSO による質心最適化、③MST による幹線抽出という三段階フレームワークで構成され、計算効率と合理性を両立する。最終的に地図上で重心・最適拠点・MST 幹線を可視化し、全国物流ネットワーク計画に資する定量的根拠を提示した。

## 4. 実験結果

### (1) データ収集と前処理

本研究ではまず、二種類のコア、データセットを構築した。第一のデータセットは、31 の省級行政区を対象とした 2024 年のマクロ経済、物流指標であり、GDP（億元）、貨物輸送量（万吨）、人口（万人）、ならびに各省都の緯度、経度座標を含む。これらの数値データは国家统计局「2024 年中国統計年鑑」に基づき、地理座標は公式測量データを参照した。第二のデータセットは、各省都間的高速道路距離、走行時間、通行料金から成り、高德地図 API により distance、duration、toll の 3 列を生成した。すべての数値単位は収集段階で統一した後、pandas を用いて両データセットを構造化 Data Frame として整理し、欠測値および異常値を検証した。マクロ指標表に欠測値は存在せず、路網表において個別エッジが欠測する場合は、後続モデルで「ユークリッド距離 × 罰則係数」により補完した。異常値は IQR 法（1.5 IQR 閾値）で検定したが、有意な外れ値は確認されなかった。さらに、尺度の影響を除去するため、scikit-learn の Standard - Scaler を用いて GDP、貨物輸送量、人口、および緯度、経度の 5 特徴を零平均、単位分散へと標準化し、後続のクラスタリングおよび最適化解析のための一貫した数値基盤を整備した。

### (2) K-means クラスタリング分析

本研究では、中国 31 省を対象に K-means クラスタリングを適用し、地理・経済的に類似したサービス区域へ区分した。指標として GDP、貨物輸送量、人口、緯度、経度の 5 変数を標準化処理した上で、 $k=2\sim30$  の範囲で WCSS とシルエット係数を算出した。その結果、Fig.2 より最適な  $k$  値の候補は 5-7 と 10-16 の二領域に分布していたが、 $k<5$  では過度の集約によりクラスタ内部の異質性が増大するため除外し、 $k=6$  および  $k=16$  の二水準を選択した。これにより戦略的マクロ区分（6 大区）と省域レベルの精密区分（16 子区）が合理的に得ら

れた。クラスタリングにより得られた中心座標は、後続の粒子群最適化（PSO）と最小生成木（MST）によるネットワーク構築の初期解として利用した。

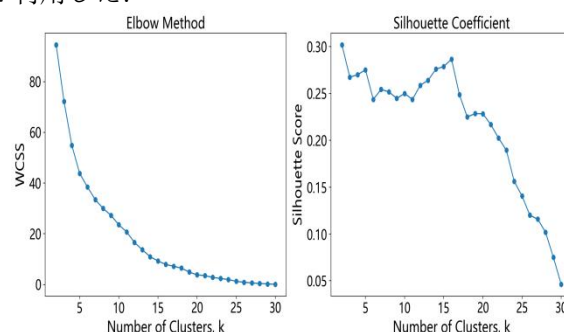


Fig.2k-means クラスタリング結果

### (3) 粒子群最適化を使用する

K-means によりクラスタが確定した後、各クラスタ内の物流拠点座標を粒子群最適化（PSO）で精緻調整した。拠点は 2 次元の決定ベクトル〈緯度、経度〉として表現し、検索範囲は当該クラスタの最小-最大緯度経度で定めた。粒子数 50、最大反復回数 100 とし、適応度関数はユークリッド距離と実高速道路走行距離、時間、通行料金の 3 成分を統合したものである（ユークリッド：実際距離重み 0.1:0.9、うち距離：時間：料金=4:3:3）。各省の貨物輸送量を重みに用いてクラスタ内コストを加重平均し、道路データが欠落するエッジにはユークリッド距離の 10 倍をペナルティとして付加した。一回 K-means + PSO：あらかじめ  $k=6$  または 16 で分割したクラスタごとに、一度だけ PSO を実行する。反復 K-means + PSO：クラスタ内 PSO により中心を更新後、再度省を再割当し、「PSO → 再割当」を中心移動量が  $<1\times10^{-4}$  となるかクラスタラベルが不変になるまで（最大 20 エポック）反復する。各クラスタのコストを集計してシステム総コストを算出し、単回方式と比較して PSO の利得を評価した。その結果、マクロレベル（ $k=6$ ）でも精密レベル（ $k=16$ ）でも、反復 PSO はシステム総コストを有意に削減し、単回方式に対して概ね 40% の低減効果を示した。これにより、後続の物流ネットワーク設計に対し、より優れた拠点座標が提供された。

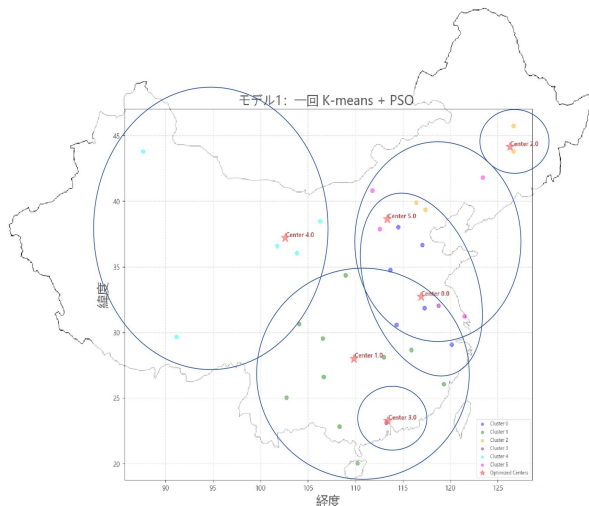


Fig.3 一回目 k-means+PSO 可視化結果(k=6)(重複がある)

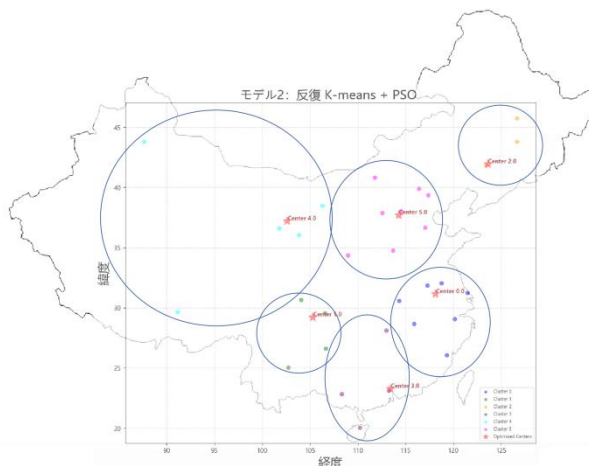


Fig.4 反復 k-means+PSO 可視化結果(k=6)(重複が改善されている)

k = 6 の実験では、単回 K-means + PSO により得られた 6 個の質心は、概ね南京、長沙、広州、北京、瀋陽、西安付近に配置された。南京、北京の質心が近接していたため、南京、合肥、鄭州、済南などの省都が二つのクラスタに分散した。長沙質心は南寧と福州を同時に管轄し、広州質心は北上して江西、湖南の一部都市を「奪取」した結果、東部 4 クラスタが大きく重複した。一方、西安質心は蘭州、銀川のみならずウルムチやラサまで単独でカバーし、サービス半径が最大となった。外層イテレーション後、質心は南京南東側、重慶と貴陽省の中間、北京の北の近郊、瀋陽、西安の東の近郊に再収束した。これにより、武漢と福州が南京クラスタへ復帰し、南寧が広州クラスタへ統合、済南と天津が北京クラスタに一本化された結果、東部 4 クラスタの境界が明確化され、重複が著しく縮小した。西安質心はほぼ不動で、依然として

蘭州、銀川、ウルムチ、ラサを担当した。調整後、東部省都から各自質心への平均円距離は約 30 % 短縮し、システム総コストは 40.51 % 低減した。これは、東中部の貨物流集積地域において、反復 PSO 機構が選址効率を顕著に向上させることを示している。

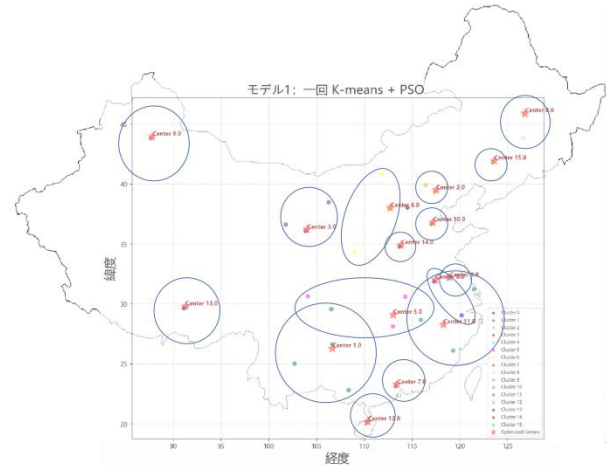


Fig.5 一回目 k-means+PSO 可視化結果(k=16)(重複がある)

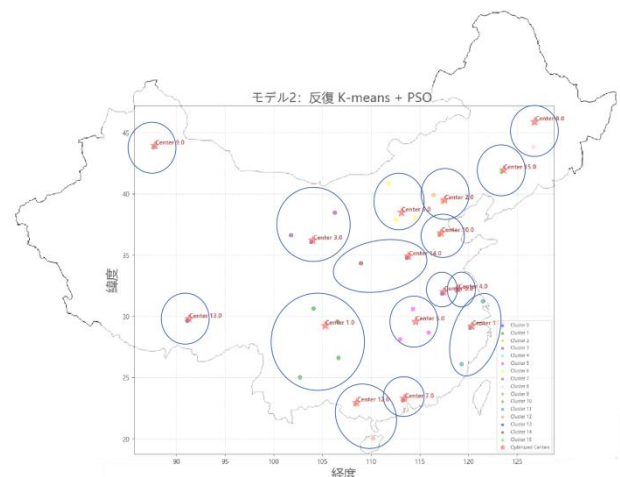


Fig.6 反復 k-means+PSO 可視化結果(k=16)(重複が改善されている)

k = 16 のクラスタリングでは、単回 K-means + PSO を適用した段階で得られた 16 個の質心は、南京、杭州、武漢、鄭州、済南、天津、瀋陽、ハルビン、広州、深セン、重慶、貴陽、ラサ、蘭州、ウルムチ、銀川近郊に位置した。

その後、外層イテレーションで反復すると、質心は貨物流集積帯に沿って再収束した。具体的には、南京質心が滬寧回廊に沿って南下し、広州質心は珠江口へさらに南移した。重慶—貴陽質心は西偏し、昆明および成都方面を統合した。済南と天津は同一質心へ統合され、福州、南昌は武漢クラスタから東南クラスタへ編入された。一方、ハルビン、ウルムチ、ラサなど周縁質心の位置はほぼ不変であった。この反復



により、東中部各省から自クラスター質心までの平均直線距離は約3割短縮され、合肥—鄭州、南寧—武漢などのクラスター間混在が解消された。その結果、単回実行方式に対して40.23%の削減効果を示した。節減の主因は、南京、杭州、済南、天津、福州など東部十余省都が最近接質心へ再マッチングされたことにある。

表1 各モデル比較

モデル	k 値	総合コスト指数(a.u.)
一回 k-means+PSO	6	247969930.42
反復 k-means+PSO	6	147522948.72
一回 k-means+PSO	16	94163652.50
反復 k-means+PSO	16	55679096.34

#### (4) 物流センターネットワーク構築

k=6の場合のMST図示は実務的意義がないため省略し、反復K-means+PSOにより抽出した16箇所の最適物流拠点をNetworkXの頂点として登録した。各拠点对の複合エッジ重みは、実道路コストと通行料に加え、道路データ欠損時の「ユークリッド距離×2000」ペナルティおよび正規化したGDP、人口、貨物輸送量に基づく経済的難易度係数を組み合わせて定義し、Kruskal法により回路を含まない最小重みの15本を選択して描画した。その結果、ウルムチ→ラサ→貴陽-重慶→南京-杭州を軸とし、北京-天津-瀋陽-ハルビンへの一本鎖が分岐する骨格構造が得られ、東中部では短距離、高密度の接続を実現しつつ、西北-青蔵地域では長距離枝を最小化して冗長投資を抑制し、16拠点の完全連結と全国規模幹線網の総コスト最小化を同時に達成した。

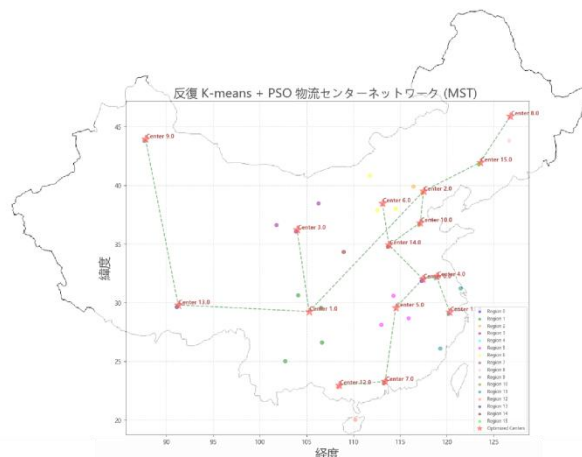


Fig.7MSTにより物流センターネットワーク可視化結果(k=16)

#### 5. まとめ

本研究は、中国31省を対象に、K-meansと粒子群最適化(PSO)を組み合わせた手法で物流拠点を選定し、輸送コストを大幅に削減しつつモデルの可解釈性を確保した点に成果がある。省域間の混在を解消し、最大で約40%の総コスト削減を実現した。さらに、得られた拠点間に対し、道路距離・時間・費用と経済・人口要素を複合重みとした最小生成木(MST)を構築することで、効率的で投資効果の高い物流ネットワークを可視化できた。結果として、東部の密集地域では短距離・経済的なルートが優先され、西北・青蔵などの低密度地域は少数の長距離幹線で連結される構造が得られた。これにより、不合理な拠点配置による資源浪費を抑え、区域経済の調和発展や持続可能な物流システム構築に寄与し、他国・他地域の物流計画にも応用可能な普遍性を示した。

#### 参考文献

- 1) 中国統計年鑑、  
<https://www.stats.gov.cn/sj/ndsj/2024/indexch.htm>.
- 2) Xinfang Zhang, Chengliang Liu, Yan Peng. Accessibility-Based Location of International Multimodal Logistics Hubs: A Case in China[J]. 2022.
- 3) SALANGAM S. 最適グローバルハブネットワークシステムの選定に関する基本研究[J]. 2017.
- 4) Dey L, Chakraborty S. Canonical PSO Based K-Means Clustering Approach for Real Datasets. Int Sch Res Notices. 2014 Nov 12;2014:414013. doi: 10.1155/2014/414013. PMID: 27355083; PMCID: PMC4897525.